FI Scientific Library

Ivan Flis

Discipline Through Method

Recent history and philosophy of scientific psychology (1950-2018)

Faculteit Bètawetenschappen FI

Discipline Through Method:

Recent history and philosophy of scientific psychology (1950-2018)

Flis, Ivan

Discipline Through Method - Recent history and philosophy of scientific psychology (1950-2018)/ I. Flis – Utrecht: Freudenthal Institute, Faculty of Science, Utrecht University / FI Scientific Library (formerly published as FIsme Scientific Library, CD-β Scientific Library) no. 100, 2018

Dissertation Utrecht University. With references. Met een samenvatting in het Nederlands.

ISBN: 978-90-73346-74-1

Cover design: Vormgeving Faculteit Bètawetenschappen

Printed by: Xerox, Utrecht

Discipline Through Method:

Recent history and philosophy of scientific psychology (1950-2018)

Discipline door methode:

Geschiedenis en filosofie van de psychologie als wetenschap, 1950-2018

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op maandag 29 oktober 2018 des middags te 12.45 uur

door

Ivan Flis

geboren op 29 oktober 1988

te Zagreb, Kroatië

Promotor: Prof. dr. L. T. G. Theunissen

Copromotor: Dr. R. Abma

For my fellow psychologists, because we need to read more.

Acknowledgments

This PhD was quite a ride from the start. In 2013, when I first found the call for project proposals at the Descartes Centre, I had no clue how far this one application would take me. I was a psychology student in Zagreb who wanted to do critical work on the scientific fundamentals of the discipline I was trained in. I contacted Ruud Abma over email to discuss the proposal. Even though he received this email from a random student from the other side of the continent, he helped me quite a bit in improving these early proposal drafts. When we finally met on August 27th, 2013 it was a sunny afternoon at Café Hofman's terrace, right after I presented the proposed research in an interview with the selection committee. The outcome was still uncertain and little did I know that the two-hour-long conversation I then had with Ruud would set the tone for the next four years. For me, this no-nonsense scholar was a real window into the critical history and philosophy of psychology. He shared so many incisive thoughts which either dismantled my half-baked ideas, or more often, made them crisp. Ruud, thank you for your erudition, criticism, and most of all friendly support from the very beginning.

Thanks to Ruud, I had the opportunity to start with and persevere in my work. Also thanks to him, I've come to grips with the 'critical' part in 'critical scholar,' the one who ended up writing this PhD dissertation. Bert Theunissen, my promotor, was a perfect counterweight to the rebelliousness I share with Ruud. He took it upon himself to reign in my irreverence and pushed me to be careful and restrained in my thinking. "Ivan, don't be so loud in your writing" he would say and usually be right. Please, don't allow for the remaining loudness in the dissertation reflect badly on him. Where the words still screech, it is my own doing. His careful reading of my work and the long discussions about the thesis and everything else an intellectual should think about were a huge inspiration and truly pushed me to my limits. Where Ruud nourished the ideas and thoughts, Bert made them shine. Bert, thanks for being both a source of intimidating criticism and invaluable support for an anxious thinker. In places where my work shines, it shines because of you and Ruud!

I received crucial support from Nees Jan van Eck at Leiden University's CWTS. Nees, thanks for recognizing the work and helping me out. Some of the most innovative aspects of what I did wouldn't have been possible without your calm intellect and knowledge.

My research took form in the day to day shared with my office mates, Noortje Jacobs, Jesper Oldenburger, Steven van der Laan, Fedde Benedictus, and later Maaneli Derakhshani. Noortje, thank you for your friendship and that big brain of yours. You've improved every bit of my Dutch experience with your caring and wit. Thanks for translating the Dutch summary, and also, for all those places where the thesis was improved by your sharp mind. Jesper, thanks for all the late evening beers, long conversations and boardgame geek-outs. They meant a lot. Steve, thanks for teaching me optimism and the importance of enjoying each day at a time. Fedde and Max, many an intriguing conversation was had during the years, and I hope, many more. Imprinted in my memory, I will carry BBG 4.07 as a special place of laughter, luminous ideas, and friendship.

The wider community of the Descartes Centre was a true intellectual home during the past four and a half years. I would like to thank Hieke Huistra in particular, for her constant support and her intellectual and emotional labor that was informal, but so precious to me. Hieke, you will always be the Source of (Caring) Authority. Here, I'd also like to thank other Descartians who I got to know and learn from: Guido Bacciagaluppi, David Baneke, Timo Bolt, Jaap Bos, Jeroen van Dongen, Friso Hoeneveld, Frank Huisman, Ingrid Kloosterman, Toine Pieters, Sandra Schruijer, Chaokang Tai, Lukas Verburgt, Rienk Vermij, and Daan Wegener. I'd also like to mention all of my students in the HPS Master's *Philosophy of Science* seminars through the years. You taught me much more than I got to teach you.

An important source of inspiration and learning were the academics I met in the annual conferences of the European Society for the History of the Human Sciences (ESHHS): Jeremy Burman, Ian Davidson, Maarten Derksen, Jannes Eshuis, Chris Green, Kim Hajek, René van Hezewijk, Jill Morawski, Annette Mülberger, Mariagrazia Proietto, Thomas Sturm, and Nadine Weidman. Two other overlapping academic groups that supported and taught me a lot was the editorial board of Shells and Pebbles and the wider Dutch community of PhD candidates in history of science. In S&P, I learned the craft of the historian of science through editing and writing, and friendly banter in our rare but important meetings with Jeroen Bouterse, Noortje Jacobs, Floor Haalboom, Sebastiaan Broere, Jorrit Smit, Hans Schouwenburg, Constance Sommerey, Emma Mojet, and Iris Clever. In the Rolduc and Glind conferences, I got to work and hang out with some brilliant budding historians like Didi van Trijp, Lejon Saarloos, Evina Steinová, Lisa Wijsen and many more. Sometime during the years, I also had the fortune to discuss psychology, science, methods, and history with an informal group of philosophers and psychologists consisting of Femke Truijens, Annemarie Kalis, Lieke Asma, Susanna Gerritse, and Noemi Schuurman. Femke, I hope there's a lot of collaborative work ahead of us!

Lastly, friends and family. First, those of you who have made life far away from home much more homely and colorful. Rodoljub and Sofija, you were a safe harbor. Jonas, Marko S., Baran, Maud, and Vedran, thank you for your company. Second, those of you who brought warmth through the screen. Ines, thanks for reading all of my work and being an awesome friend taking care of me from back home. Anja, Filip, Lorena, Chris, Lorenz, Munish, Iva, Kosta, Selim, Panos, Antonia, Marko D. ditto on those closest of friendships temporarily (or perpetually?) at a distance. You guys kept my sanity.

Dubravko, hvala ti za svu podršku kroz sve ove godine. Ja sam možda ludi brat s glavom u oblacima, ali ne bi ta glava tamo bila bez tvog držanja "lojtri." Tata, mama, hvala vam za sav trud, rad i ljubav koji ste uložili u svoje sinove. Ove stotine stranica koje slijede nikada ne bi bile ispisane bez vas. Možda vam to nisam uvijek znao reći na dobar način, ali iskru svega što sam u životu uspio napraviti dobio sam kroz vašu ljubav i potporu. Negdje sam još uvijek ono dijete s legićima, samo su legići danas uglavnom riječi.

En fin – Megi. Bila si uz mene gotovo pola mog života, tijekom svakog uspona i pada. U ovu disertaciju utkana je sva tvoja pažnja, požrtvovnost, ljubav i strpljenje, ali i tvoj prkos i nepoštivanje autoriteta. Ironija toga da je tvoja potpora dovela do dugog i kompliciranog teksta, baš onakvog kakve uglavnom mrziš, mi nije promakla. Dapače, to je baš jedna od onih malih smješnih stvari koje se nakupljaju i čine naš zajednički život krasnim. Hvala ti što me činiš boljom osobom svaki dan. P.S. hvala majčici za sve vožnje na aerodrom i što te napravila takvu kakva jesi.

"If a book were written all in numbers, it would be true. It would be just. Nothing said in words ever came out quite even. Things in words got twisted and ran together, instead of staying straight and fitting together. But underneath the words, at the center, like the center of the Square, it all came out even. Everything could change, yet nothing would be lost. If you saw the numbers you could see that, the balance, the pattern. You saw the foundations of the world. And they were solid."

Ursula K. Le Guin, *The Dispossessed*

"The scientist is like the gambler, the businessman, the physician, or the psychotherapist. He knows in advance that no matter what policy, guideline, or rule of thumb he follows, and no matter how clever he is in implementing it, he is sampling from the vast universe of facts and he is working much of the time in the semidarkness of more or less ignorance. Hence all guidelines are inherently stochastic, both as to their descriptive and prescriptive roles."

Paul E. Meehl, *Cliometric Metatheory*

Chapter 1. Introduction

In February of 2017, Chris Noone had defended his thesis and received a PhD in psychology. Chris is an Irish psychologist who finished his Bachelor's in Galway, continued his studies with a Master's in applied neuroscience at Leiden University in the Netherlands, and then returned to Galway to finish his graduate research. For three years, he conducted research in order to produce evidence about the possible effect of mindfulness training on critical thinking skills. By producing this kind of evidence, and writing it up in a lengthy thesis, he also qualified for receiving a PhD. To qualify for a PhD, Chris conducted three studies. The first used a cross-sectional individual differences design, the second an experimental manipulation, and the third a randomized-controlled trial. These are very specific names for particular research practices of psychologists. While setting up his studies and interpreting the results, he took great care in canvassing the scientific literature on mindfulness and critical thinking to inform his judgment, and considering the results and the background information, he concluded that (Noone, 2016, p. V): "The results of these studies together suggest that the effects of mindfulness on critical thinking are mostly small and, in experimental contexts, indistinguishable from those of closely matched control conditions." In other words, Chris's expert judgment informed by his research was that mindfulness training will not help much with critical thinking.

Across the continent, on the other side of the European Union, Kosta Bovan was also handing in his thesis. Kosta is a Croatian psychologist who finished his Bachelor's and Master's in psychology at the University of Zagreb, and then continued his graduate research at the same university. He set out to investigate the concept of voting in democratic elections: What does it mean for a voter to vote "correctly" in elections? And how can scientists experimentally investigate voting behavior? Kosta also designed three studies. In the first, he created mock newspapers' headings and topics of interest in order to prepare an artificial political campaign for the participants of his experiment. With the second study, he calibrated the questionnaire in relevant individual differences the participants exhibited, and the third as the centerpiece of his approach: A quasi-experiment that put the participants into an artificial election campaign and asked them to gather information in a carefully constructed environment and vote based on that information. What Kosta did was also an implementation of the research practices of psychologists, mixed in with the approach from the kind of political science being done at the University of Zagreb. His aim was investigating elections through the lens of psychologists' current ideas about information processing and decision-making. The results of his study showed that "those participants that had better political knowledge" had been "exposed to lower cognitive load and used complex strategies of decision" (Bovan, 2016, abstract) and consequently, had a higher probability of casting a correct vote.

Chris Noone and Kosta Bovan are just two psychologists who concluded their graduate research in psychology in 2016. Worldwide, hundreds of thousands of others have done so by participating in a very particular way that psychologists think about problems and answer practical and theoretical questions. Should we invest in mindfulness training of students to increase their critical thinking capacities? Is there a causal link between being mindful and the cognitive functions that are used when we think critically? How should modern liberal democracies ensure that their citizens are capable of making informed choices when voting? Which theory of rationality should we accept, considering the evidence about people's strategies when making their voting decisions? These are just some of the questions Chris and Kosta wished to answer with their research. The questions this PhD thesis will attempt to elaborate is the way psychologists in recent history go about asking and answering scientifically interesting questions. The motivation behind the thesis is trying to understand how psychologists set up their research problems, canvass the scientific literature, devise research procedures, choose what statistical analyses to use, and argue for their conclusions.

I am interested in the way psychologists peer into the world. Can we describe that way of peering into the world of human psychology as distinct from other approaches, and if yes, where did it come from and where is it going? In the rest of the introduction, especially in section 1.2, I will use specific examples from Kosta's and Chris' theses to showcase the development of certain features of psychology as a scientific discipline in recent history. Think of it as a reading tool bringing closer the convoluted historical and philosophical questions that are the topic of this thesis. The reading tool serves the purpose of bringing those issues to a more day-to-day level – so quotidian for psychologists that students pursuing a doctorate need to master them.

I zoomed in on the research of two people in particular as a contrast to the approach I will take in this thesis. My approach is general, non-local, and long-term. **General** in the sense that I am not interested just in one or two areas of research, but the most broadly constructed area of research that is recognized as scientific psychology in the second part of the 20th century. **Non-local** in the sense that I am not interested in what happened in Galway, Leiden, Zagreb, or Cambridge on this or that side of the Atlantic – but most places that housed scientific psychologists producing research in recent history. **Long-term** in the sense that I am not interested in what happened in psychology in 2016 or 1966, but in roughly the last seventy years, from the end of World War II to the first two decades of the 21st century.

Even a cursory consideration of the problems related to general views on psychology leads to a truism: Talking about scientific disciplines as a single thing is a daunting task. In that regard psychology is no exception. Roger Smith, in the *Fontana History of the Human Sciences* which condenses "what historians have written, discussed, and debated about psychology's history over the last 40 years" (Stam, 2013, p. 2), puts it bluntly: "*There is no one discipline of psychology*" (Smith, 2013, p. 14, emphasis in the original). Smith calls psychology a "family name for a bewildering range of beliefs and occupations" (p. 12). Indeed, the mass of practices – be they practical or academic – is staggering and unsurveyable.

Saying "Psychology is…" without qualifications will lead to simplification, distortion, and possibly reification of perspectives that are neither relevant for psychologists, nor for other educated audiences living in the 20th and 21st centuries psychologized societies.[1] And yet, this thesis will be an attempt at a general perspective, qualified, however, by the historical and philosophical research that has been done on 20th century psychology. Thus, in the methodological sense of talking about science historically or philosophically or sociologically, my argument is in part about the value of general perspectives *in general*. Taking a bird's-eye view on things is a valuable activity in itself, especially when it powers reflection and criticism.[2] The usefulness of generalist discussions of psychology, and especially historical arguments, will be a reoccurring topic in the following chapters.

The general perspective in the thesis will be constructed out of three things – simple and down-to-earth content of introductions to psychology in undergrad-level university textbooks in the period from 1950 to 1999 (Chapter 2 & 3); the specialized technical discourse that can be data-mined from abstracts and titles of articles in psychology journals from 1950 to 1999 (Chapter 4); and the highly volatile discussions about what is good science in the 2010s debates centered on the replication crisis in psychology (Chapter 5).

The view of disciplinary history I will retrieve from the above has a few inbuilt limitations: It is **internal**, because it is produced by psychologists. It is **in English**, because it is produced by English-speaking psychologists, be they native or non-native speaking groups of academics participating in 20th century scientific Anglophonia.[3] It is heavily **biased toward American psychology**, because many of the actors worked at American universities, and even when they did not, they more or less participated in American intellectual traditions of late 20th century psychology. And, finally, it is **discursive**, because it is constructed out of different kinds of linguistic representations of psychology.

In the rest of the introduction I will attempt to qualify my perspective on psychology's disciplinary formation. In section 1.1, I will contextualize the connection between scientific psychology as a global phenomenon in late 20th century and the formative

[1] For an elaboration on 20th century societies becoming psychological, see Smith (2013, p. 102-136), Jansz and Van Drunen (2003), and especially the work of Nikolas Rose (1990; 1998).

[2] For a more elaborated polemical argument arguing that historians have neglected more general perspectives (what historians like to call *longue durée*), see Guldi and Armitage (2014); and the commentary of historians of science in an *Isis* Focus Section (Jacobs, 2016).

[3] For a thorough historical study of the role of language in natural science and the rise of global scientific English in the twentieth century, see Gordin (2015).

influence American psychology had in that development. Here, I will also include a very brief prehistory of the scientific psychology that will be the topic of this thesis, by discussing the formative decades of late 19th and early 20th centuries. With section 1.1, I have two aims. The first is to ground my work in the historiography of psychology. The second is to specify what I mean by twentieth century scientific psychology. In section 1.2, I will discuss the different elements of that scientific psychology in detail: operationism, inferential statistics, construct validity theory, the expanding scientific literature, and psychologists' genre of writing. The aim of this section is to more thoroughly inform my view with the historical and philosophical research that was done so far. In section 1.3, I will discuss psychology's replication crisis which is taking form in the 2010s. I use the replication crisis as a magnifying glass that explicates the elements of scientific psychology in the words of the psychologists themselves. In section 1.4, I will shortly use the qualified perspective on psychology as a discipline I developed in the introduction to structure and introduce the chapters of the thesis and define my research question.

1.1 Grounding scientific psychology in 20th century US

The historian Adrian Brock (2006) provided three tongue in cheek rules of inclusion/exclusion in the history of psychology, telling us whose work we might find when leafing through books that include chapters on the discipline's history:

> Rule #1: If your work did not have a major impact on American psychology, however influential it might have been elsewhere, it does not count.
>
> Rule #2: If your work had a major impact on American psychology, even though its influence was limited or nonexistent elsewhere, it is an important part of the history of psychology.
>
> Rule #3: Asia, Africa, Latin America, and Oceania do not exist.

Brock's sarcastic rules reveal a basic intuition held by contemporary psychologists and historians of psychology – psychology as an academic discipline is a thoroughly American affair. As the historian Wade Pickren puts it: "With the economic and military ascendancy of the US after World War Two, this thoroughly American psychology was exported around the globe. It became redundant to use the modifier, American, before psychology" (2009; p. 87). If you were participating in *any* line of research identified as scientific psychology in late 20th century global north (and even beyond), your thinking and practices must have been exposed or even thoroughly informed by American psychology. Not all psychology between 1950 and 2018 is

American psychology, but the academic parts that aren't, are constructed in direct comparison or opposition to this multifaceted tradition.[4]

Both Brock and Pickren are historians, and they are arguing against this domination of American psychology – either because it is a distortion produced by those parochial Americans who are uninformed about the rest of the world, or, more accurately and less sardonically, because it is the product of a systematic kind of one-sidedness in perspective that has exacerbated throughout the 20th century. The aggravation of one-sidedness happened both in the academic histories of the discipline and the self-image of its disciples. By a meandering and contingent process of cultural, political, scientific, and military dominance of the United States in worldly affairs, and an unprecedented amount of funding for research at American universities,[5] post-WWII psychology and the history of that psychology worldwide were thoroughly Americanized.

To historians, the extent of this Americanization became evident when their mode of analysis moved from seeing psychology as an abstract application of the universal scientific method toward the view that psychological investigative and professional practices are a sort of social activity. The business of producing psychological knowledge and plying the psychologists' trade, with its developing norms and institutions, and most importantly, with its history; was thoroughly social.[6] In light of the work of historians of psychology, taking the view that there is just one psychology became untenable. If psychological knowledge was produced in locations in time and place, and that localization framed what kind of knowledge was being produced and how, the idea that there is one scientific psychology became self-contradictory.

An insightful parallel to this crumbling of the narrative of one (American) psychology is comparing it to the view of science largely shared by contemporary historians of science. In the past decades, they have stopped talking about the scientific method as an abstract universal and instead started amassing empirical studies of change-through-time of scientists/natural philosophers' manifold practices, discourses, and

---

[4] A great example of how pervasive American psychology was globally in the late 20th century is the history of psychology in the ex-Yugoslav states. Kosta Bovan, the political psychologist I mentioned earlier, works in a thoroughly American tradition of scientific psychology, despite the fact his *alma mater* was a university in a communist state during the Cold War. How this tradition came to exist at universities in ex-Yugoslav states is still not historically investigated, despite being an extremely interesting question for social history of science.

[5] For a detailed study of the role of funding in the American social and behavioral sciences, see Solovey (2013). For the role of the National Institute of Mental Health in psychology in particular, see Pickren and Schneider (2005).

[6] This is the perspective advanced by historians of psychology at the end of the 20th century. For some of the most influential work advancing this view, see Danziger (1990; 1997), Morawski (1988), and Smith (2007; 2013).

objects of study. As Daston and Galison put it when discussing scientific objectivity as a type of epistemic virtue: "[E]pistemic virtues do not annihilate one another like rival armies" but "rather they accumulate" (2010, p. 363). They accumulate in historical time, in a social context, and are mediated by the thinking, talking, and writing of specific individuals organized in groups.

If we keep the perspective of historians of science in mind, the move of historians of psychology from viewing one tradition as dominant in the abstract toward a plurality of perspectives in particular can be seen as just one disciplinary history entering the hallowed halls of contemporary academic history of science. In line with this crumbling of meta-narratives, some critical psychologists and historians of psychology seem to argue that psychology itself, and not only our view of its history, is truly becoming local – a patchwork connecting different indigenous psychologies around the world, the American just being one of many.[7]

1.1.1 Ignoring the elephant in the room

The big narrative of scientific psychology might have lost its promise of a total description, but I would argue that we should not ignore the fact that it is still powerful, in the sense of its sheer size and soft power.[8] Among all the disconnected pieces that make the mosaic of twentieth century psychology, some of them do form a comprehensive federation of thinking and acting "styles."[9] Some psychologies, in the variety of research programs of the past decades, did coalesce and form more compact wholes – a great example for this is how the great number of psychotherapies developed during the 20th century were scientifically reined in by standardized psychotherapy efficacy research (Wampold, 2001). In turn (or maybe because they could stick together in such a way), they were received by historical actors – be they psychologists or other experts – as more scientific than others. They didn't only hold

---

7 Arguments for the idea that we are heading for a more inclusive kind of psychology and history of psychology, as to the types of descriptions of subjectivity that are seen as scientific, are many. They include the previously mentioned Danziger (1994; 2006), Brock (2006; 2014), and Pickren (2009), but also critical historians/psychologists like Thomas Teo (2006; 2015). Brock, however, recognizes some problems with the development in which all historians of psychology just become historians of science studying psychology (see Brock, 1995).

8 I take the notion of power in the analysis of history of psychology from Nikolas Rose (2006, p. 106-107): "Power in the case of psychology would not be thought of in negative or instrumental terms, as that which manipulates, denies, serves other purposes. Rather, psychology would be viewed from the perspective of the 'power effects' that it has made possible. For psychology, like the other 'human' sciences, has played a fundamental role in the creation of the kind of present in which we in 'the West' have come to live."

9 I say federation, because I am reluctant to call it a system – as many of the arguments in this thesis will show, scientific psychology is in part internally inconsistent and incoherent. It is more a federation of outlooks of like-minded actors than a formalized meta-narrative. For "styles" of scientific thinking or reasoning as an analytical category in the history and philosophy of science, see Hacking (2002; p. 178-199).

an epistemic higher ground, for whatever reason, but they also had clout because of their size and efficient expansion.

Whether the normative epistemological proclamations of those clumps, which were recognized as more scientific, should be accepted or rejected is open for critical debate. Indeed, it has been frequently debated by methodologists, historians, philosophers, and sociologists. Psychologists themselves have given these discussions many names through the decades: introspection vs. observation of behavior, idiosyncratic vs. nomothetic, clinical vs. scientific, and qualitative vs. quantitative just being a few of the more prominent ones that were used as rallying cries for reform of the whole field, with more or less success.

The questioning of the epistemological justification of psychologists' research translated into a destabilizing influence. Critical discussions about the scientific authority of psychologists lend themselves to a state of permanent instability.[10] However, despite the internal and external anxiety about psychologists' epistemic authority, we shouldn't lose sight of the fact that a circumscribed notion of scientific psychology did arise from the myriad debates during the 20th century. This is what I mean by saying that different practices of scientific psychology have stuck together to form a mass that has its own pull on psychologists and their cultural context. In other words, different ways of doing psychology in recent history – and, by extension, the psychological knowledge those practices produced – enjoy the imprimatur of societal and academic legitimacy depending on to what degree they conform to the kinds of psychology recognized as scientific.[11] In the highly complex psychologized societies of the global north in recent history, the epistemological norms of that loose set of scientific psychology pervade the state and society through healthcare, education, entertainment, and management. Scientific legitimacy does not only provide intellectual gains but is also the coin of the realm in Western democracies in the late twentieth century.

Kurt Danziger gave this connection between the ability to productively employ your knowledge in society and epistemological legitimacy the name "marketable methods," and it works both ways – the methods structuring society and the society selecting what kinds of methods flourish. As he put it (1990, p. 101):

---

[10] For an analysis of how psychologists from the 19th century onward keep declaring 'crises' of their discipline in order to advance certain kinds of psychological thought, see Sturm and Mülberger (2012).

[11] Keep in mind that my argument does not suggest a single cause or lineage for this scientific psychology. Scientific psychology is a product of a complex set of historical contingencies. It is definitely not the end-point of a straight line of development. Likewise, I would caution the reader against reading my description as an invocation of scientific progress. Even when I say scientific psychology expanded, I am agnostic about its progress.

> The fact is that almost from the beginning of the twentieth century psychology ceased to be a purely academic discipline and began to market its products in the outside world. That meant that the requirements of its potential market were able to influence the direction in which psychology's investigative practices were likely to develop. Practices that were useful in the construction of specific marketable products were likely to receive a boost, whereas practices that lacked this capacity were henceforth placed under a handicap.

The psychology indigenous to American universities in early and middle twentieth century – the collection of practices so well described by Kurt Danziger and historians of experimental psychology in the United States (Morawski, 1988) – was the historical context that produced the conventions of scientific psychology this thesis will study. Scientific psychology's far reach and influence in late twentieth century are a consequence of, among other things, the way it was abstracted and isolated from its historical context and social reality. This abstraction and decontextualization turned it into a scientific psychology that spread worldwide during the decades of the Cold War and early 21st century.

The basic practices of American scientific psychology – experimental and correlational research designs, inferential statistics, operationalization of variables, measurement and test theory, the production of a massive and constantly growing literature, and the scientific article as a medium of communication – started living a life of their own and defining what scientific psychology is for hundreds of thousands of students, professionals, and scientists around the world.

It took decades for the conventions of scientific psychology to spread in research and teaching. The medium of that expansion was their institutionalization – the most prominent examples being textbooks, guidelines for writing journal articles, graduate school requirements, and technical recommendations and handbooks issued by professional organizations. On one hand, the wider institutionalization of these conventions entrenched the asymmetry of the center-periphery relationship between different kinds of psychology done in the global north and the rest of the world. They defined the kind of arguments and the kinds of vocabularies different psychologists could employ while still remaining comprehensible to each other. On the other hand, the conventions also drew borders within global northern psychologies. They formed a scientific pecking order that intellectually and institutionally pushed the psychologies that would not conform to the periphery.

While the conventions of scientific psychology provided a common ground for debate, psychologists pursuing varied programs with different philosophical underpinnings could choose to more or less align their way of doing psychology with these conventions and stand to gain something from that alignment. Every such move came with rewards or setbacks, depending on local conditions and the direction of the move.[12] If groups or individuals conformed to the conventions well, or manipulated them in an innovative way, their ideas had the chance to circulate far enough and break through Brock's three rules. If not, they stood almost no chance to be heard in the cacophony of psychology in the second part of the twentieth century.

The conventions, deeply entrenched in the "epistemological unconscious" (Steinmetz, 2005) of the discipline, started becoming more and more invisible, as accepted customs are wont. They would usually resurface as elephants in the room at times of controversy. At calmer times, the conventions remained the background that was easily ignored, a kind of implicit consensus that is not seriously discussed or available to criticism.

1.1.2 Intellectual and physical geography

Up to this point, I have interchangeably called the scientific psychology I am describing American, Western, and global northern. The most precise label would be Occidental or global northern, because by the beginning of the 21st century, centers of scientific psychology are spread out throughout the northern hemisphere.[13] A great example for this far reach of scientific psychology is the recent replication crisis – the growing controversy is not only affecting American departments of psychology, on the contrary, it is in its nature international (more on this in section 1.3).

Locating scientific psychology is only partially an issue of physical geography – the center-periphery relationship also works for some traditions of psychological thinking in the global north. The convoluted history of psychoanalysis and its interaction with academic psychology is a stark example of this, with psychoanalysis and scientific psychology growing either closer or more apart in different places and times, depending on the local conditions and traditions. Psychoanalysis and scientific psychology jostling for epistemological legitimacy could be narrated as one of the central debates in the 20th century history of psychology, and in fact, it has been by some scholars (Jahoda, 1977).[14] Similar histories, albeit less pervasive, could be told for

[12] Sometimes literally with a reward. See Sandra Schruijer's (2008; 2012) research on the American funding of Western European experimental social psychology during the Cold War. Even when prominent experimental social psychologists tried to establish a distinctly European tradition of experimental social psychology, they ended up using American funding to launch research programs that conformed with the conventions of scientific psychology I am describing in this thesis. Schruijer asks the critical question whether, in light of that American influence, we can actually say that there is a distinctly European tradition of experimental social psychology after the Cold War?

[13] For the terms "intellectual geography" and a discussion of "center-periphery" applied to history of psychology, see Danziger (2006). Danziger's polycentric yet asymmetric account of the development of scientific psychology up to the middle of the 20th century informs much of the view recounted in this subsection.

[14] Jahoda's account is particularly interesting when she contrasts what she calls academic psychology to psychoanalysis. Then she discusses questions that could be directly quoted into the main body of this thesis,

overtly politically positions on human psychology, like Marxism or feminism. This is also why I decided to call my object of investigation recent scientific psychology, and not recent American psychology. Not all American psychologies of the 20th century fully, or some even partially, conform to the basic elements of the scientific psychology I am describing in this thesis.

Returning to the physical geography of scientific psychology, a thorough study on how abstracted international trends impacted a single national context in the global north is Trudy Dehue's (1995) work on the changes in Dutch psychology during the twentieth century. Dehue's (1995, p. 150) argument is that the "empirical-analytical methodology [what I call scientific psychology] could prevail […] in the 1950s and 1960s" in the Netherlands because of the changing relationships between workers and employers in Dutch society. These changes in society were then "sustained and contributed to" by "modernization-bound psychologists," whom I would call scientific in this thesis. The Dutch methodologists epitomized by the likes of De Groot were producing a distinct psychological tradition that could be compared to what quantitative psychologists were doing in the United States, but, as Dehue argues, these American and Dutch scientific psychologies were not equivalent. In the Dutch case, they elaborated the kind of psychology that could answer the niche "opportunity structures"[15] framed by Dutch society – Dutch scientific psychologists adapted well to the changes in larger society and then reproduced them in their science.

Dehue's Dutch case study provides an insightful example of how scientific psychology in the 20th century was a loose collection of family resemblances, and not a close-knit fixed system. However, my emphasis is different than Dehue's, because I stress the similarities of slightly different psychologies identified as scientific in the past seventy years, while Dehue emphasizes their differences and incoherence. I argue that there is a set of conventions of scientific psychology that coalesced during the 20th century, while she stresses that no such thing can survive closer historical inspection. As she puts it: "the position of international agreement on empirical-analytical rules fails to hold water. On closer philosophical consideration most Dutch psychologists will turn out to follow a quite individual variation within the international empirical analytical-framework" (Dehue, 1995, p. 128). This is precisely where I diverge with some historicist perspectives on 20th century psychology. My argument is that the focus on local contexts very often occludes broader trends. One of the broader trends that became invisible is the main topic of this thesis, namely, the broad trend of the inertia of scientific psychology.

like: "[D]o available methods, sanctioned by their similarity to natural science procedures, define the problems which are legitimate research topics in psychology, or do legitimate problems require the inventions of new methods appropriate to them?" (Jahoda, 1977, p. 153).

[15] For the concept of "opportunity structures" within history of science, see Huistra and Jacobs (in press).

1.1.3 Inertia

Identifying the indigenous systems of thought that run counter to the dominant one, historically and in contemporary science/societies, is not only intellectually satisfying, but also politically liberating for the practitioners of those non-mainstream ways. Liberating for particular national styles of psychological thought, like psychology in Africa,[16] but also ones in the global north that are explicitly political like feminist psychology.[17] Postcolonial, subaltern, critical, feminist, indigenous, non-universal perspectives exist and speaking of *the* scientific psychology will not do, historians and critical psychologists have argued. While agreeing with historians like Danziger that what is psychology is multifarious, my overarching argument in the thesis is that in that diversity a set of conventions of scientific psychology have amassed a sort of inertia that is glossed over.

What I mean by intellectual inertia is that scientific psychology remained fundamentally unchanged through time, despite shallow breaks and local idiosyncrasies, or even damning criticism. As it formed in the decades after World War II, scientific psychology expanded in size and dearth, but its main tenets and direction remained the same. The institutionalized conventions reproduced and expanded from one community of psychologists doing research to the other, who then started producing scientific literatures as prescribed and made possible by those conventions. Conventions and literatures of scientific psychology moved like massive intellectual glaciers, grinding unconventional psychologies into obscurity or making them conform and stick to the glacier's mass.

In view of this inertia of scientific psychology, the pluralist perspective of historians and critical psychologists needs to be checked not because it is descriptively incorrect, but because it is incomplete as it downplays the amassed inertia of a kind of abstracted and generalized scientific psychology. It focuses on the pebbles that have survived the

---

[16] For attempts at constituting African psychologies that are distinct from global northern psychology, see Nwoye (2015) and Ratele (2017).

[17] On the history of gendered epistemology in American psychology, see Morawski & Agronick (1991). However, keep in mind that even the psychological research motivated by progressive politics like feminism does not by definition exclude what I am describing as scientific psychology. On the contrary, some feminist scholars use the authority of scientific psychology in the way described by Eagly and colleagues (2012, p. 212) in their overview of feminist topics of research in the late 20th century: "In most of the research that we discuss, authors have not explicitly addressed this gender equality goal, nor have they labeled their research as feminist. Nonetheless, the gender-equality goals of feminism have no doubt led many researchers to investigate topics such as sexism, sexual harassment, and violence against women that implicitly or explicitly relate to feminist goals. Such value-directed choices do not invalidate the research, given that all scientific research stands or falls according to the replicability of its findings and the critical scrutiny of communities of researchers." In other words, feminism informs the motive, science informs the knowledge produced. A cynical reading might be that scientific psychology accepts feminism, as long as the feminists don't criticize epistemology.

grinding, or even whole mountain ranges chained in ice, but doesn't say much about the glacier looming large. Coming back from the glaciological metaphor, the fact that we can inspect and describe all those different kinds of psychology does not mean their multitude is changing the universal. Put differently, although we know of many groups of "alternative" psychologists and psychologies, this doesn't mean that the dominant tradition has disappeared or diminished. Universal narratives do not lose their causal efficacy or power because we see them as socially constructed. Scholarly description does not negate power asymmetries, neither in the epistemological nor in the political sense.

Meta-narratives and capillary soft power being abstract concepts, let's discuss it by means of an example: The usual textbook American historiography of scientific psychology from behaviorism to the cognitive revolution. Kurt Danziger gives historians and psychologists a good and important scolding for seeing that historical shift as anything more than changes limited to American psychology. Danziger argues that the cognitive revolution is falsely seen as an episode of global proportions because of American hegemony:

> Major themes in the American context, like behaviorism, are relegated to minor footnotes, and other themes, unknown to most American psychologists, become highly significant. Important developments for American psychology, like the cognitive revolution, turn out to be non-events from a European perspective, because of the existence of a local cognitivist tradition that never managed to cross the Atlantic. (Danziger, 1994, p. 476-477)

For Danziger, both history of psychology and psychology as a scientific discipline are leaving a sort of Dark Age of American parochialism. He optimistically says in the same text: "[M]odern psychology is returning to the position from which it began: a polycentric position in which there are diverse but intercommunicating centers of psychological work" (1994, p. 477). I would caution against such optimism. Despite the fact that history of psychology has gone through its renaissance of a new approach since the 1980s, the perspectives developing among some historians and critical psychologists have still remained marginal in psychology at large.[18] In the above example, the idea that the cognitive revolution opened the black box of the mind and liberated psychologists' thinking – the thinking of *all* psychologists – is still alive and kicking as a productive fiction in the wider discipline. It is a productive fiction in the sense that it provides one historical episode for grounding a professional identity.

[18] Probably the most ambitious example showcasing the limit of historically informed thinking influencing psychology is Kenneth Gergen's (1973) proposal for "social psychology as history." Despite being popular among some critical social psychologists and historians, it wasn't taken up by the wider discipline.

The reason why the importance of the cognitive revolution has remained a productive fiction for psychologists at large has nothing to do with norms for historical research, but it has everything to do with the inertia of that fiction. For a small group of historians of psychology, scientific methodologies of 20th century American psychology *are* socially constrained and historically contingent practices. Danziger, in particular, played a huge role in conducting research on the social history of methods, with his descriptions and critical discussions of the “rise of the aggregate” (Danziger, 1990), “methodological imperative” (Danziger, 1985), and the history of variables (Danziger & Dzinas, 1997). Despite the work of historians like Danziger, for most other scientific communities of psychologists in recent history, methods are not a product of social history. They are black-boxed[19] givens carried through time on the shoulders of giants.

1.1.4 Prehistory – from the late 19th century

The cognitive revolution as a development relevant for psychologists at large is just one topic in the productive fiction of scientific psychology in recent history – the kind of historical narrative that justifies it well after its beginning. The disparate parts of the beginning of that tradition, the one I call recent scientific psychology, can be identified in the way that the problem of quantification of psychological phenomena was resolved in the early twentieth century. From the debates among the scholars in psychophysics and the mental testing movement, who formed distinct communities of psychologists and philosophers largely based in Germany, the United Kingdom, France, and the United States, a consensus and a set of practices emerged that ignored the difficult philosophical questions about the nature of psychological phenomena and whether they could be quantified.[20] As Gail Hornstein (1988, p. 2, emphasis in the original) put it:

> During the period roughly spanning the years 1860-1940, two distinct transformations took place. First, a wide range of psychological phenomena (for example, perception, intelligence, personality, and learning) were fundamentally *redefined* to make their basic properties quantifiable. Second,

[19] ‘Black-boxing’ is a term commonly used in science studies: “An expression from the sociology of science that refers to the way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become” (Latour, 1999b; p. 304).

[20] The history of how early 20th century psychologists methodologically resolved the issue of quantification of psychological phenomena from within psychophysics and mental testing was written by Gail Hornstein (1988) and I extensively draw from it in this section. Hornstein, in turn, draws from a large body of historical scholarship on 19th/early 20th century experimental psychology and mental testing. For a more thorough view, also see Danziger (1990) and Rose (2006). For more up to date work, see Araujo (2016) on Wundt and Carson (2007) on mental testing.

> psychological phenomena that resisted quantitative treatment (for example, forms of consciousness, spirituality, and will) were *jettisoned* from the domain of legitimate empirical psychology. As a consequence, psychology shifted in the American context from being a discipline which seemed inherently non-quantitative to one that relied almost exclusively on quantitative approaches.

She (Hornstein, 1988, p. 3) also notes that these negotiations are usually "omitted from the standard histories" which in turn "renders scientific approaches static and ahistorical." This is because vital programs of research need to tell their histories as series of progressive steps of success and crucially so when it comes to their own empirical fundamentals.

Hornstein also gives two examples (1988, p. 7 & p. 12) of how "ignoring the opposition" was a successful strategy for founding research traditions. "Ignoring the opposition" is the name Hornstein gives to the early 20th century psychologists' practice of ignoring fundamental criticism of their approach because the criticism could not be accommodated in any satisfying way. 19th century and early 20th century psychophysicists went on measuring despite the fact that what and how they were measuring was questioned. A similar thing happened with measuring intelligence in the 1930s – the fact that no consensus existed over what intelligence was, had become a non-issue because the testers could point toward how practically useful and widely used intelligence testing was.

In the following section, I will briefly cover similar resolutions of later contentious issues for scientific psychology in the second part of the 20th century – the longwinded debate over operationism and the one about significance testing. I will also cover other elements of scientific psychology that have their usually obscure histories and philosophical implications, and whose institutionalization through conventions created the inertia of scientific psychology in recent history.

1.2 Elements of inertia of scientific psychology

In this section, I will describe a number of elements of scientific psychology. My aim is to more precisely identify what I mean by scientific psychology and provide an argument as to why this complicated system of ideas and practices has intellectual inertia.

By the end of the 20th century, the "deep" work of justifying scientific psychology's practices, goals, and norms was either settled, or the unsettled controversies were systematically forgotten. The conventions of scientific psychology were elaborated and expanded during the second part of the 20th century by hundreds of thousands of researchers worldwide. As I will argue, scientific psychology was a set of conventions which researchers followed, and by doing so, produced a massive literature that was in practice unsurveyable.

Contemporary historians of psychology may be uncomfortable with such a general statement of chronological constancy that imputes a commonality to a varied and huge collection of actors. This is why I will review the scholarship of historians (and historically-minded philosophers) themselves on a few crucial topics in the period: operationism, inferential statistics, psychological constructs, the size of psychology's academic literature, and psychologists' genre of writing. My aim is to provide a sketch of the conventions that sustain scientific psychology – not a perfect closed system specifying a foundational methodological and epistemological program, but a messy and vague collection that lumbered through the decades in a relatively stable, but never compact, form. In this part, I will also make use of Chis Noone's and Kosta Bovan's theses, to exemplify the elements of scientific psychology's inertia in the actual practices of contemporary psychologists.

1.2.1 Operationism

In his thesis on mindfulness and critical thinking from 2016, Chris Noone dedicates two sections of his introduction to discussing "operational definitions" of mindfulness that psychologists use. Before mindfulness – as a practice, but even more importantly, as a knowable concept – can enter into scientific psychology, it needs to have agreed upon operational definitions. Operational definitions are a kind of translation from terms we use in everyday speech into the repertoire of scientific psychologists. In order for something to be even considered as a subject of interest, it needs to be rendered into an object suitable for scientific psychology. Such objects, when operationally defined, live strange lives. Some concepts have multiple competing operational definitions. Some definitions are used for multiple concepts. There are still other operational definitions that do not latch onto any everyday psychological concept at all, because they are actually used as translations between different expert jargons.

While canvassing the psychologists' literature on mindfulness, Chris has identified a number of operational renderings of mindfulness by psychologists. He needed to carefully consider them and argue for his own operational definition, which he then manipulated and dissected experimentally to produce evidence about: "Since the middle of the last decade, several operational definitions [of mindfulness] have been proposed, each influenced by different traditions within psychological science. Common to many of these definitions of mindfulness is a focus on self-regulation" (Noone, 2016, p. 11-12). However, there was no consensus: "Attempts at operationalising and measuring mindfulness have not been any more successful in producing consensus" (Noone, 2016, p. 19) The notion of operational definitions is one of crucial elements of recent scientific psychology and it has an intriguing past.

The history of operationism spans the twentieth century. The position was originally proposed by the physicist Percy Bridgman in the 1920s and 1930s (1927; 1938), and then elaborated by psychologists Stevens (1935a; 1935b) and Tolman (1936).[21] The basic idea of operationism was that a concept of scientific interest, like Chris' mindfulness, was synonymous with the kind of operations scientists would use to measure and manipulate it. This view never fared that well within physics, but it sparked a long list of debates and positions in twentieth century psychology. Through that time, a philosophically loose form of operationism also became institutionalized as a convention of scientific psychology.

Uljana Feest (2005, p. 131), in her conceptual and historical analysis of operationism, gives the following outline of the spurts of the debate among psychologists and philosophers: After the first elaboration by Stevens and Tolman, operationism was debated in a number of papers of *Psychological Review*. This string of papers culminated in a symposium on operationism (discussed in detail by Green, 1992b). Feest then identifies the debate and the refinement of the concept through the 1950s, then again in the 1980s, 1990s, and the 2000s.

In the early days of the discussion, in the context of behaviorism, the important question about operationism was its relationship to either logical positivism or different strands of neobehaviorism. Later, after World War II, the discussion became less epistemological and more methodological. By contrasting the epistemological and methodological aspects of the debate, I am following Feest's (2005, p. 134) distinction that "the aim of epistemology is to provide a theory of what it would take to justify existing systems of knowledge", whereas methodology is a brand of epistemological thought that tries to "formulate guidelines for the acquisition of new knowledge."

What can be identified, both from Feest's and Chris Green's (1992b) analysis, is that different historical actors took operationist positions for different reasons. What is also evident is that their operationist positions differed. The philosophical side of the debate was about the ramifications of adopting operationism as a form of a neutral observation language. However, psychologists did not apply operational definitions in research practice in order to justify their knowledge epistemologically, but to have a system that formalized and organized "their prior assumptions about how to interpret observations" (Feest, 2005, p. 145).

In other words, providing stripped down operational definitions allowed them to render their production and interpretation of data as a comprehensible system that was public. Other psychologists could inspect and debate it in the literature, and then add to it or change it. Justification of knowledge was the business of philosophers,

[21] For the historical account of how operationism arose as an interdisciplinary type of "scientific philosophy" elaborated in the loose interdisciplinary culture of mid-century Harvard, see Isaac (2012). Isaac also investigates in detail the role of Stevens, Boring, and Skinner in bringing operationism into psychology (2012, p. 92-124). Another important historical actor in Isaac's "scientific philosophy" at Harvard was the philosopher W.V.O. Quine, whose work will be briefly discussed in Chapter 5.

while psychologists occupied themselves with the nitty-gritty of making it. They were the hard-nosed experimenters producing data, not armchair scholars trying to make sense of it conceptually.

Part of Feest's argument is that the fact that operationism was tied to quantitative approaches in psychology is a historical accident – it might as well have been thoroughly qualitative. For her, this is not a question of quantitative versus qualitative psychology, but, more fundamentally, of the viability of empirical psychology: "To put it quite bluntly, *all empirical psychologists have to operationalize their concepts*. And all empirical psychologists have to argue for their results by laying open both their conceptual presuppositions and their empirical data. Viewed this way, the issue seems to be whether psychology can be an empirical science at all" (Feest, 2005, p. 145). Indeed, all empirical psychologists must render their thinking into something that can be observed in some way. Doing it through quantified operational definitions, though, also allows for something of crucial relevance for mid-20th century scientific psychology: Operational definitions give psychologists the kind of raw input for the machinery of inferential statistics to draw conclusions from their data.

Operational definitions and inferential statistics, for psychologists, connect observations and the theoretical thinking about observations. Considering this, at the time of the discussions of Tolman and Stevens, the connection between quantitative and operational might have been a matter of a historical contingency. However, following the institutionalization of inferential statistics in psychology after World War II, those two things became inextricably linked. The connection is corroborated by the fact that non-quantitative approaches were effectively marginalized in the recent history of scientific psychology. The number of non-quantitative articles in the psychological literature since the 1950s is extremely small, especially outside of the specialized journals that cater to qualitative researchers specifically.[22] In the view I am developing in this thesis, the varied approaches conventionally grouped under the name of psychology's qualitative methodologies are marginalized by definition, because they do not conform closely enough to the conventions of scientific psychology.

Feest and Green agree that what operationism could not do is "ultimately replace the fruits of hard, rigorous thought" (Green, 1992b, p. 315). I would claim operationism was not even competing with conceptual analysis in post-WWII scientific psychology, because analysis of generalizations and theories was almost fully exhausted by inferential statistics. Quantified operational definitions just represented the kind of

[22] For a detailed study on the representation of qualitative research in Anglophone psychological journals since the 1950s, see Marchel and Owens (2007).

input well suited for the statistical procedures of inference-making that promised much vaunted objectivity through numbers.[23]

1.2.2 The institutionalization of inferential statistics

Inferential statistics in social science have a rocky history, and particularly so in psychology. Inferential statistics are the kind of statistical analyses that are used to draw conclusions either from samples about populations or as mathematically formalized explications of reasoning about hypotheses. From a very marginal role in the 1930s, a kind of inferential statistics became the “sine qua non of scientific inference” by the 1960s (Gigerenzer et al., 1990; p. 210). To reason scientifically became virtually coextensive with reasoning statistically. Psychologists, like other social and medical scientists, have imported a particular way of statistically testing hypotheses from population genetics in the form of Fisher’s testing method in the 1930s and 1940s. By the 1950s, Fisherian methods were graduate program requirements at American universities despite the fact that a fundamental criticism of Fisher’s approach existed since the 1920s, which was elaborated into a different theory of inference by Jerzy Neyman and Egon Pearson.

Instead of adapting one or the other, or rejecting both; psychologists have “tried to fuse the controversial ideas into some hybrid statistical theory” (p. 208; Gigerenzer et al., 1990; hereafter I call it NHST, short for **N**ull **H**ypothesis **S**ignificance **T**esting).[24] This hybrid statistics was then institutionalized across the board of sub-disciplines in psychology - as a pedagogical requirement for students, as journal policy, and as a normative standard for scientific inference. The institutionalization was differently received in different sub-disciplines and research groups, and the different trajectories of the adoption of significance testing would make for an expansive study of its own, but what is mostly agreed upon is that significance testing of some kind has become

---

[23] For the now standard historical work on the connection between numbers and objectivity, see Porter (1995). For the long history of objectivity in scientific thinking, see Daston and Galison (2010).

[24] For a review of NHST debates, see Nickerson (2000). Nickerson also includes a very interesting note in the beginning of his article reviewing the NHST controversy: “One of the people who gave me very useful feedback on a draft of this article questioned the accuracy of my claim that NHST is very controversial. ‘I think the impression that NHST is very controversial comes from focusing on the collection of articles you review—the product of a batch of authors arguing with each other and rarely even glancing at actual researchers outside the circle except to lament how little the researchers seem to benefit from all the sage advice being aimed by the debaters at both sides of almost every issue.’ The implication seems to be that the ‘controversy’ is largely a manufactured one, of interest primarily—if not only—to those relatively few authors who benefit from keeping it alive. I must admit that this comment, from a psychologist for whom I have the highest esteem, gave me some pause about the wisdom of investing more time and effort in this article. I am convinced, however, that the controversy is real enough and that it deserves more attention from users of NHST than it has received” (2000, p. 241). It goes to show how efficiently ignored oppositions can be institutionalized, so much so that they are made invisible even fifty years after entering the scene.

the inferential standard if one did quantitative work in psychological research in the second part of the twentieth century.

On the flip-side, this institutionalization of the hybrid inferential statistics did not go without opposition. Since NHST was first used there was a protracted discussion about what the proper inferential test was, and how to best employ it in psychological research. Some connect the beginning of the opposition to when psychologists got caught in the crossfire between the Fisher and Neyman-Pearson testing method in the first decades of the twentieth century (Acree, 1978). Others claim it has been part and parcel to the century long discussions over statistical inference (Hacking, 1965/2016). In psychology, we can trace it already to 1960, at the very least, to a paper by Rozeboom (1960, p. 416) who called it the "dogma of inferential procedure which, for psychologists at least, has attained the status of a religious conviction."

The discussion, at first, boils down to the question: Can psychologists reject a hypothesis based on a statistical test? What it is actually about for the most part is: Which statistical test should psychologists use to reject hypotheses? The distinction between these two questions is important. The former questions the norm itself, the latter its technical implementation.[25] The latter one is actually asking what kind of statistics is appropriate if researchers want to infer something from their datasets. What is particularly interesting, especially for historians, is the length of the discussion - it has been going on for at least sixty years, with no settlement in sight - and also the polemical, and sometimes quite brutal, tone of it.

Here are a few examples of the prolonged debate's extreme tone from its later episodes in the 1990s. The first example is a paper by Jacob Cohen (1994), who cynically entitled his criticism of NHST abusers and dogmatists, as he saw them, as *The Earth is Round ($p < .05$).* Cohen's witticism deftly exposes the anxieties over the malleable and easily manipulable standards of evidence psychologists have institutionalized during the 20th century. The psychologist Gerd Gigerenzer (2004), in a now famous paper under the biting title *Mindless statistics,* accused social scientists of substituting careful inspection of data for ritual when drawing conclusions. In the 1990s and 2000s, he published a number of papers on what he calls the "null ritual" that hinders scientific progress (Gigerenzer, 1998b; Gigerenzer, Krauss, and Vitouch, 2004).

The length of psychologists' discussions over significance testing can be seen as a parallel to the protracted debate Uljana Feest and Chris Green identify for operationism. Both NHST and operationism were seen as epistemologically controversial by a minority, and at the same time, both were institutionalized across

[25] If one just questions the technical implementation of inferential statistics, it is taken for granted that the psychologists' object of research is well represented by numerical aggregates in the first place. Kurt Danziger offers a critical discussion about this kind of "triumph of the aggregate" (Danziger, 1990, p. 68-87). I will briefly discuss this in the next subsection, through the work of Joel Michell.

the board by simply "ignoring the opposition" à la the debates Hornstein (1988) described in psychophysics and mental testing. The trench war over NHST, and the declaration of crisis of the standards of inference, was confined to particular circles of psychologists and methodologists. The inferential standard of hybrid statistics mostly reigned freely and supremely outside of those circles, both in textbooks and in peer reviewed journals. Gigerenzer and colleagues (1990) call this the "inferential revolution."

What this revolution also entailed, besides a shift in the statistics and the arguing for hypotheses, was a shift in the design of experiments. As they put it: "Fisher has linked significance testing to experimental design, and the 'inference revolution' was consequently a revolution in experimental design…[that] has been so successful that it is often difficult for today's experimenters to imagine that 'experiment' could mean something different from what Fisher had taught" (Gigerenzer et al., 1990). The hybridized procedure that arose out of the merging of the Fisher and the Neyman-Pearson procedure kept this quality of retroactively structuring experimental designs; and I would go even further to say that it also structured non-experimental designs that employ quantitative methods.[26]

To sum up, in the middle of the twentieth century a very specific and constricted procedure for making inferences had developed and spread through the communities of psychologists. The statistical procedure got equated with the scientific approach almost to the exclusion of all others, and it also structured the already quantified methodologies of research that were in place (the experimental design, but also what is called correlational research – this distinction will be the topic of Chapter 4). This approach was institutionalized through education, journal policy, and implicit values of good science - and alongside its institutionalization, there was constant and very outspoken criticism.

Kosta Bovan, in his thesis on voting behavior, also employed inferential statistics.[27] For the many potential variables connected with the concept of accurate voting under investigation, he performed various inferential tests to confirm whether the data gathered from his participants exhibited statistically significant patterns of response. Depending on their significance or lack thereof, he compared his results to the ones

[26] A great example for the structuring of experimental designs is the rise of randomized controlled trials as the gold standard for research on efficacy of psychotherapy (for a critical perspective, see Truijens, 2016). The history of randomized controlled trials in psychology is more complicated than just NHST, which only reinforces the belabored point that the conventions of scientific psychology are the result of a contingent historical process. There are many moving parts. For a critical take on the role of NHST in more non-experimental research, see Meehl (1990b).

[27] NHST is the inferential standard taught today at all Croatian psychology departments on the undergraduate and graduate level. For a critical contemporary study of how it is institutionalized through education in psychology, see Flis (2013; note that the Master thesis is written in Croatian).

found in the literature – agreeing with some, disagreeing with others, and calling for more research on most. Theoretical work on the relationship between various constructs[28] he was working with in his thesis was facilitated by employing inferential statistics, and based on the results of those tests, specifying which relationships between constructs are robust and which should be inspected more closely.

Evidently, the institutionalization of this brand of inferential statistics was also a perfect fit for the practice of producing operational definitions. As a psychologist, you had a theoretical construct like mindfulness as in Chris Noone's thesis. You then provided a set of operational definitions that made that construct into something that could be linked to a set observations. You set up your studies in such a way that they yielded observations that were relevant for the particular operationalizations of your construct; you collected the data and produced aggregates as numeric formalizations of the operational definitions; and you tested that aggregate with the inferential statistics that are officially sanctioned by journals you aim to publish in.

This procedure leads you to a couple of options: Accept the construct as defined, refine it, or reject it. Using inferential statistics and operationalizations in this way, however, required you to utilize a formalized system for conceptual analysis that rendered "constructs" into something intelligible. Operational definitions and significance testing, however powerful in turning the psychologists' object of research into something manipulable, did not give you an abstracted level of generalizations one would expect from scientific work. Talk of constructs did that, and we will turn to it in the next subsection.

1.2.3 History and philosophy of psychological constructs[29]

The institutionalized "inference revolution" from the middle of the twentieth century provided a decontextualized and universal procedure for drawing conclusions about hypotheses based on empirical data. It spread the notion that in order for psychologists to test a claim based on their empirical studies, they needed to corroborate their conclusion by some statistical test of significance. However, a different kind of usage of statistics was also becoming widespread for scientific psychologists in the period, and that was the descriptive kind – statistics that are used to represent the objects of scientific interest numerically. I already shortly discussed it based on the history written by Gail Hornstein, and I will expand on that now, casting my look forward toward the second part of the 20th century.

---

28 'Construct' is a technical term used by psychologists that I will discuss in detail in the following subsection.

29 This section is largely based on the work of Kathleen Slaney (2017) and Joel Michell (1999). Slaney's book is the most thorough philosophical, historical, and practical evaluation of construct validity theory. Michell published multiple papers and a book on the history and philosophy of measurement in psychology.

Important and highly critical historical and philosophical work on the history of psychologists' measurement – how psychologists render their objects of scientific interest quantitatively – was done by Joel Michell (1999). Michell's (1999; 1997) argument is that psychologists since the middle of the 20th century have almost universally accepted the definition of measurement as "formulated by S. S. Stevens" in 1946 which is that "measurement is the assignment of numerals to objects or events according to rule" (Michell, 1997; p. 360). Michell finds Stevens' definition of measurement astonishing (the same Stevens we encountered in Green's and Feest's discussion of operationism) because of the simple fact that this is not the definition of measurement seen in the natural sciences, the scientific older sibling psychologists supposedly take inspiration from. In the natural sciences, measurement is usually defined as establishing ratios through experiments, or "the process of discovering or estimating the measure of some magnitude of a quantitative attribute relative to a given unit (Michell, 1999, p. 14.).

The "assigning numerals according to rules" definition of measurement did not remain an obscure quirk promulgated by Stevens, but as Michell argues based on his inspection of books in the period from 1950 to 1999, it spread far and wide among psychologists. "[P]sychology, as a discipline, has its own definition of measurement, a definition quite unlike the traditional concept used in the physical sciences" (Michell, 1997, p. 360). Stevens extended his definition of measurement practically, giving psychologists the simple system of defining measurement based on different scales: nominal, ordinal, interval, and the ratio scale (see Michell, 1986). The distinction between the different scales became part of the methodological furniture – a basic element, as I call it – of scientific psychology.

The combination of operationism and Stevens' definition of measurement were just a part of the far-reaching developments in psychological measurement that turned into a decontextualized convention of scientific psychology. The other are the so called test theories – classical test theory that was developed up to the 1960s, and the modern test theory that succeeded it afterwards.[30] Test theories are big container names for psychometric descriptions of measurement devices (usually tests, thus the name) and the data produced when using them. The central concept of the classical test theory was the "true score model, according to which an individual's observed test score is conceptualized as being composed of two non-overlapping parts: a 'true score' and a 'error' component" (Slaney, 2017, p. 32). The innovation in the modern test theory that followed it was "that the interitem structure of a set of test data may be represented well by one or more latent variable models" (Slaney, 2017, p. 32). The practical ramifications of these test theories for empirical research and theory construction in 20th century psychology were mediated and channeled through the specification and development of construct validity theory.

[30] For a technical treatment of test theory, see McDonald (1999).

Slaney broadly describes construct validity theory as "a general theoretical approach and set of methods for judging whether empirical inferences and decisions made on the basis of quantitative data are licensed by the most current theory regarding the 'construct' purportedly measured by the test or assessment tool in question" (2017, p. 1). The basic outline for construct validity theory was set in the report *Technical Recommendations for Psychological Tests and Diagnostic Techniques* issued in 1954 by a number of organizations, chief among them the American Psychological Association. This report was republished and changed in the following decades, and it represents the kinds of standards that tried to specify and suggest best practices for psychologists doing measurement. What stayed constant was its focus on validity, and to a certain extent, its focus on constructs.

This first report, from 1954, was finessed in a hugely influential paper by Lee Cronbach and Paul Meehl, *Construct Validity in Psychology Tests* (1955). Elaborating on construct validity has since then become a cottage industry of "validity theorists," scientists who were elaborating, criticizing, and expanding construct validity in the following decades. Their elaboration collectively went in the direction that, both the idea of validity and of constructs, "have become much [*sic*] increasingly more flexible and, thus, more vague" (Slaney, 2017, p. 135).

Slaney provides an informed historical argument that construct validity theory, from its inception in the 1950s, channeled a few seemingly disconnected debates psychologists were having up to that point. Namely, the ideas of classical test theory and latent variable models with their issues of validity and reliability, whose basic outline was laid out by Charles Spearman, and the neobehaviorists' discussions about intervening variables and hypothetical constructs. The talk of validity and factor analytically produced latent variables like *g* (general factor of intelligence) was then systematized into construct validity theory in the 1950s.

The systematized position produced standards for conducting research with psychological tests, and most importantly, validating them. During the next seventy years, it also produced a riot of epistemological positions.[31] Some psychologists perceived constructs as ontological objects – the real entities they were after when conducting their research. Others perceived them as useful handles for organizing data collection, instruments used in thinking about psychological concepts. Whether instrumentalist or realist, all of them had to account for the relationship between constructs, their operationalizations, and data; considering that: "Construct validation is involved whenever a test is to be interpreted as a measure of some attribute or quality which is not 'operationally defined'" (Cronbach & Meehl, 1955, p. 282).[32]

[31] See Slaney (2017, p. 143-236) for an overview and main actors and their positions.

[32] Cronbach's role in developing construct validity theory was also expressed in his idea that psychology is separated into two large communities, "correlationists" and "experimentalists." Construct validity theory was, for Cronbach, a way for unifying those two methodological traditions "as a method for theory-building

Validating constructs, thus, became a required practice when working with psychological phenomena.

Whatever the highbrow philosophical reading of construct validity theory was, its function has been in bridging the gap between the extremely constrained operational definitions and the broader theoretically interesting concepts behind them. For philosophically-interested psychologists like Meehl, the bridging was among other things justificatory because it provided scientific legitimacy to psychologists who produced research by using operationalizations of constructs.[33] But construct validity theory's even more far-reaching effect was that it practically rendered possible data production about theoretically interesting objects of investigation. With its minutely bureaucratic ways of specifying how to talk about constructs in relation to data, construct validity theory provided psychologists with a language for theorizing. Not only for the philosophically-minded like Meehl, who debated some of the most sophisticated philosophers of science of the time, but also for the thousands of "grunts" of psychological research in all the trenches of applied psychology and their varied departments that existed since the 1950s. In other words, psychology got a theoretical language that had nothing to do with compromised theoretical positions like behaviorism or psychoanalysis, and psychologists got a science with a method and a way for elaborating theory, without any specific theory in mind - an open-ended unfinished science suited for the empirically-minded scientists who wanted to dispense with talk of theory and metaphysics.

Psychologists' effort to develop construct validity theory in order to grasp the theoretically interesting things behind operational definitions is also a direct corroboration of Uljana Feest's claim that psychologists, at least after the 1950s, *did not* embrace operationism with the intention to "exhaustively provide the meanings of the concepts in question" (2005, p. 143). Your average psychologist did not accept construct validity theory and operationalization for the purpose of adhering to some philosophical theory of meaning, but for pragmatic reasons, so they could produce data and theoretical interpretations about that data. Most importantly, this allowed communities of psychologists in late 20th century to systematically expand their science to a wide range of topics. This, in turn, facilitated the deluge of psychological research spanning every facet of Westerners' daily life I describe in our analysis of psychology's literature in Chapter 4.

Construct validity theory, as it was developed since the middle of the 20th century, was a framework for systematically connecting the way psychologists measure things, the

in psychology generally" (Slaney, 2017; p. 102). The failure of Cronbach's unifying project will be discussed in detail in Chapter 4.

[33] How this aspect of construct validity theory factors into the psychologists' intuition about the way that their science builds and tests theories will be discussed at length in Chapter 5, in the context of the kinds of philosophy of science that has traction among psychologists in the replication crisis debates.

way they make inferences about the things measured, and the way they talk about them theoretically. Theoretical discussion of phenomena, for scientific psychology, was thus exhausted by the formalized conceptualization and refinement of constructs facilitated by inferential statistics. Slaney (2017, p. 210) comments on the conceptual confusion that arose from talk of constructs, considering that constructs have become both "theoretical concepts" that are "created and/or used by researchers to designate and communicate about the phenomena under study" and, at the same time, the actual phenomena they are referring too. As Slaney continues to argue, they are perceived as unobservable both when used as theoretical concepts and as referents for actual phenomena. When used as theoretical handles, constructs are unobservable because they are talked about in the abstract. When used as referents for actual phenomena, they are again unobservable because the phenomena are only accessible through measurement instruments. Seen like that, scientific psychology is a way of producing evidence about unobservables, because the talk of constructs makes psychological phenomena empirically unreachable in principle.

The content of scientific psychology after the 1950s, as I understand it in this thesis, is an amalgamation of descriptions of constructs. As Slaney (2017, p. 202-203) puts it:

> [...] [T]he concept 'construct' has been, and continues to be, used to denote a very large class of phenomena in psychological and related discourses, including more classically defined traits (such as introversion and extroversion), clinical and diagnostic categories (e.g., psychopathy), cognitive functions (e.g., cognitive control, verbal memory), and more specific attitudinal and/ or behavioral phenomena (ranging from "attitudes towards work schedules" to "pharmacists' care of migraineurs").

Psychologists were in the business of elaborating the operationalizations of already existing constructs, coming up with new ones, refining old ones, and making the most marketable among them practically relevant for uses in society at large. They had in place a uniform methodological system for investigating phenomena institutionalized through statistical routines and bureaucratic procedures, which allowed them to divide themselves into separate-yet-loosely-connected communities organized around different sets of constructs and still remain within the same scientific discipline. An internal division of labor, if you will.

Walter Mischel comically called the extreme version of this division of labor the toothbrush problem: "Psychologists treat other peoples' theories like toothbrushes – no self-respecting person wants to use anyone else's" (Mischel, 2008, para. 3). A set of constructs could be pursued as far as there was agreement and room to publish about it. If things changed for the worse for a particular construct, one could just move on to a new set, by either developing her own from scratch or by modifying older ones. This was not only acceptable, but even desirable, because it showed how psychology allowed for "scientific progress." Discussions around constructs were organized into different subdisciplines of psychology, and the written form of these discussions of constructs was a massive, ever-growing literature. The crucial role of that massive literature will be discussed in the following subsection.

1.2.4 The massive literature

It is a matter of demographic and scientometric fact that psychology as a profession and as a science enormously expanded during the 20th century. James Capshew (1999, p. 1) in his history of American psychologists' role in the world wars, listed the staggering figures of the number of psychologists in the United States: "Between 1919 and 1939 the number of psychologists grew tenfold, from approximately three hundred to three thousand professionals. [...] The psychology community had expanded by another order of magnitude by 1970, with more than thirty thousand professionals registered as members of the American Psychological Association. In 1995, fifty years after the end of World War II, the number of psychologists in the United States was approaching a quarter of a million." It wasn't only the number of individual psychologists that grew in orders of magnitude – their scientific output followed suit.

At least in the rate of expansion of its literature, psychology is no different than the natural sciences. Already in 1963, Derek de Solla Price had argued that each field of science approximately doubled its literature every ten to fifteen years (De Solla Price, 1963/1986). The data on the size of psychology's literature in Chapter 4 conforms to this growth pattern of the natural sciences – psychology's literature, as represented by American Psychological Association's database *PsycINFO,* doubled in size in every successive decade from 1950 to 1999. The steady growth of the literature is also strongly related to language considering that most of the journals indexed in *PsycINFO* are published in English.

In the continuously expanding landscape of psychological literature of the 20th century, American psychology is not only the largest literature; it is also the largest one that can be accessed in an organized way. The structuring and organizing of psychological literature also has a long 20th century history, with the first bibliometric aid called *The Psychological Index* developed in 1895, then replaced by an intricate abstracting service of *Psychological Abstracts* in 1927, which during its 80 years of existence slowly transformed into the digitized modern database of *PsycINFO.*[34]

The digital transformation of how psychology's literature was organized on this meta-level also conveys a basic fact about the shift in how scientists go about accessing published articles – from pursuing the printed copies of individual journals earlier in the 20th century, to journals being organized into databases by the end of the century. Instead of looking into a particular journal, a scientist perusing the literature searches different databases by using keywords, especially when trying to get a comprehensive

[34] For a detailed chronology, see Benjamin and VandenBos (2006).

picture about research on a certain topic. The search algorithms retrieve individual articles from potentially thousands of journals, forming a list of bits and pieces that is tailored for answering the search query of the scientist.

PsycINFO, and *Psychological Abstracts* before it, can't be seen as just bibliometric tools. They are also a type of information and social technology developed in order to parse and structure psychology's protean literature. They make it comprehensible for humans. Instead of hundreds of thousands of disconnected articles about varied constructs, they give it a semblance of structure. Jeremy Burman makes a similar argument when identifying possible National Science Foundation sources of funding for the development of *PsycINFO*, connecting it to "Vannevar Bush's 'memex proposals': the imagined construction of a collective memory machine" (Burman, 2018, p. 21). Burman especially focuses on the internally developed 'controlled vocabulary', a system of keywords which employees of *PsycINFO* assign to individual articles. He goes even further, calling for a Foucauldian kind of historical analysis into the relationship between government funding, bibliometric aids, and the content of psychological science. Power and knowledge are not linked only through social practices, but also information technology.

Until such a Foucauldian reading of *PsycINFO* is developed, I would argue for a more modest view: There is a strong reciprocal connection between the expanding literature of English-language psychology and the conventions of scientific psychology as I described them thus far. Each reinforced the other. Conducting research as outlined by operationism, institutionalized significance testing, and construct validity theory, psychologists could pursue and expand productive research programs on numerous topics. A powerful metaphor for this view of psychology's literature in the period is a nomological network, the idea that scientific theories are a coherent system of theoretical constructs and observations connected to each other. The view of the nomological network was very important for the debates over construct validity and its connection to logical empiricism (see Part II in Slaney, 2017). Psychologists expressing the view that they are building a coherent nomological network by adding studies to the literature will be described in a number of chapters of this thesis and I will come back to it in the conclusion.

For now, the important point is that psychologists could multiply research studies in a systematic way and have community-wide debates about collections of such studies organized around particular constructs. They could also apply this mode of research to anything that humans in recent history recognized as psychological or related to human or animal behavior. In turn, such a mode of arguing about psychological phenomena would rapidly grow in size and become an expansive model of research for other psychologists. By its size and scientific credentials, the literature would become not only large, but also prestigious and a normative showcase of *the* scientific psychology.

Kosta Bovan, our newly-minted political psychologist from Zagreb, also gives an odd hint at this in the acknowledgments of his thesis. Among the usual group of family and friends one thanks when finishing the arduous work of writing a PhD thesis, Kosta also thanks a woman called Alexandra Elbakyan "for offering graduate students across the world the opportunity to finish their studies." (Bovan, 2016, acknowledgments). Alexandra Elbakyan is the controversial founder of the largest pirated database of scientific literature in history, called *Sci-Hub*.[35] Kosta, as a student in a scientifically peripheral country like Croatia,[36] works at a university with extremely slim funding for university libraries so his access to the global Anglophone academic literature of journals and academic books is constrained. Access to literature is a rudimentary requirement for even passively participating in global scientific psychology, let alone the active participation Kosta is aiming for.

The size of psychology's literature wasn't only a source of pride and structure, it was also a cause of concern for 20th century psychologists. A peculiar kind of technique for both surveying and expanding the burgeoning literature was the development of a writing standard and a specific genre that psychologists employ when writing research reports.

1.2.5 Codified genre of scientific writing

Different scientists in different periods write in different ways. This truism becomes obvious when we acknowledge the fact that scientific knowledge is produced by groups of people in specific socio-historical contexts. To produce knowledge – to make claims about the world – one needs to put it into a coherent stream of thought that is comprehensible to others: First to the community of like-minded thinkers, and then to a potentially hostile (or just uninterested) humanity at large. Alan Gross (1990, p. 3) puts this succinctly: "Rhetorically, the creation of knowledge is a task beginning with self-persuasion and ending with the persuasions of others." The form that scientific writing takes largely depends on the community of scientists that write and read it.

A 20th century high-particle physicist discussing bubble chambers and a 19th century alienist talking about *dementia praecox* wrote in very different ways. This kind of discrepancy gives us an opportunity to look at the history of scientific communities as a rhetorical history of the different genres they produced. Charles Bazerman (1987; 1988) investigated this kind of history of scientific rhetoric by looking at psychologists' papers as texts conforming to a particular genre.

Bazerman focused on the way psychologists' writing developed around the American Psychological Association's Publication Manual which codifies the so-called APA

---

[35] For more on Sci-Hub, see Himmelstein and colleagues (2018) and Bohannon (2016).

[36] For a comprehensive study of Croatian scientific production in the social sciences and the humanities at the end of the 20th century, and why the country is classified as scientifically peripheral, see Jokić, Zauder, and Letina (2012; written in Croatian). For a less comprehensive source in English, see Letina (2016).

style. The style's official history harks back to 1929, when an APA internal committee produced a seven-page guide for authors submitting their papers to the journal *Psychological Bulletin*. We will come back to the specific history of those guidelines at the end of this subsection. For now, it is important to know that by 1952, these guidelines turned into an official publication manual for the psychologists' community at large, and the manual went through a number of editions and is still published today. It became widely used not only by psychologists, but other social scientists as well.

In a more philosophical view, Bazerman contends that since the manual developed during the height of behaviorism in American psychology, it codified a sort of behaviorist rhetoric. By examining over a hundred articles in the period from the end of the 19^{th} century to 1980, he constructed a chronology of changes in the psychologists' empirical report genre, which I will shortly discuss. The changes in the genre, according to Bazerman, mapped onto the shifts in thinking about psychological topics.

In the earliest period up to the 1920s, the articles mixed philosophical and experimental exposition – they didn't have a uniform system of subheadings nor a clear separation into different sections. They were written as "continuously reasoned arguments" (Bazerman, 1987, p. 268). In the interwar period and after WWII, the reports became more structured and uniform along the lines of what Bazerman calls behaviorist rhetoric: "The previous tendency toward low-level conclusions that give only aggregate descriptions of the behavior observed no longer is a difficulty - it is the whole extent of the enterprise. One looks only for patterns of behavior, not underlying principles or mental operations" (1987, p. 272). Result sections became dominated by statistical talk, figures, and tables. The author of the text was not "a reasoner about the mind," s/he was rather a "doer of experiments, maker of calculations, and presenter of results" (1987, p. 272). Method sections decreased in size and elaboration, because they were not places in which the authors introduced innovations. Instead, method sections were used as justifications for the conclusions and interpretations that were being made in the text. The research report stopped being a reasoned argument and turned into a disjointed collection of elements labeled by standardized headings: "The results become the core of the article. Discussion often merely sums up the data and is sometimes relegated to small print. Conclusions do little more than repeat confirmation of the descriptive hypotheses" (1987, p. 272).

The most interesting conclusion Bazerman draws from his analysis is that the new genre was designed in such a way that it facilitated the addition of new, albeit small, pieces to the psychologists' project of description of behavior. Bazerman calls this "incremental encyclopedism" (1987, p. 273). He lists several rhetorical consequences of viewing psychology as an exercise in incremental encylcopedism. First, statements of the research hypotheses became crucial, moved to the front, and were repeated multiple times throughout the text. Second, since psychologists were adding to a comprehensive project, articles shrunk – less work was needed to justify methods and looking at the researched phenomena from many angles. Third, the newly adopted referencing style clearly marked the surname of the author and the date of the publication, giving a visual cue to the incrementalism. Authors and the timestamps of their contribution were mustered as corroboration of claims (more on this in Chapter 2). Fourth, the reader was not approached anymore as a person trying to understand or solve some problem. They were assumed to be looking for pieces of understanding to fit into their system of thinking, and to look for faults why the bits in the particular research report might not have been produced in a sound way. A basic consensus about topics and objects of research was implied, the author just had to persuade the reader that s/he was arguing about those topics and objects in a sound way.

Bazerman's elements of the genre stand even for the period after the 1980s. If we take a look at Chris Noone's and Kosta Bovan's theses, we will find the characteristics of the genre Bazerman described. In Chris' case, his chapters are structured like individual articles *because* they are published as articles. His thesis, much like this one, is a collection of published articles or articles under review. Those articles were written in the style roughly conforming to the genre Bazerman describes. Kosta's thesis was written as a monograph, but even in the case of such a more integrated work, we can recognize the familiar structure in the way he discusses his research through a method, results, and a discussion section. Both theses conform to the genre prescribed by the APA manual, both in their structure and in the referencing style.

The more problematic part of Bazerman's conclusion is identifying the genre with behaviorism. Matthew Sigal and Michael Pettit (2012) have persuasively argued that the formalized style wasn't developed to project behavioristic thinking onto wider psychology, but with a much more pragmatic goal in mind: to cope with information overload. The uniform style, as discussed and developed in the form of the guidelines by the APA committee in 1929, had the aim of making "the burgeoning journals more financially secure and to render the psychologist's reading practices more efficient" (2012, p. 362). A uniform style was a writing technology, invented to help deal with the rapid increase in the number of relevant publications, which was a constant problem for editors, publishers, and psychologists as consumers of journal articles. The situation from which the guidelines arose in 1929 is also a great example of how a very particular cultural context – in this case of 1920s American push for industrial rationalization – produced an abstraction that spread through institutionalization far and wide in the following decades. Among hundreds of thousands of others, in 2016, it also reached an Irish psychologist investigating mindfulness and a Croatian one looking into voting behavior.

I would agree that the consequence of the new genre, with or without intention, was the kind of efficient incremental encyclopedism Bazerman described in his painstaking reconstruction of psychologists' rhetoric. The genesis of that encyclopedism was less philosophical and more pragmatic but its influence was far-reaching. Incremental encyclopedism was just exacerbated by the end of the century because the primary organizational structure through which psychologists accessed literature moved from the individual journal, to a database producing strings of individual articles when prompted by a search query.

The article format[37] that became a standard in psychological literature meshed perfectly with the other previously described elements of scientific psychology. At her disposal, the researcher in any of the many psychology's subdisciplines from mid-20th century onward had a standardized system to elaborate her theories in the nascent and developing construct validity theory, a way to connect those constructs to empirical claims through operationalization, a set of agreed upon procedures for statistically testing those quantitative operationalizations for significance, and an economic and smart format in which to report all those things. Following those conventions for analysis/writing/thinking, studies could be produced on a potentially endless number of topics, in potentially endless variations that added to the incremental encyclopedism describing behavior and human psychology.

Psychology, in any shape and form that would minimally conform to the institutionalized conventions I have described thus far, had a language, a method, and the idea that if enough studies were produced, a comprehensive theory would arise out of the accumulated research reports. How this system reproduced itself through the work of hundreds of thousands of psychologists until today is best revealed in the contemporary heated debates centered on the replication crisis. The replication crisis reveals and exposes the fault lines of scientific psychology, because of some unforeseen interactions between the conventions. Psychologists could voraciously expand their discipline and its subdisciplines, but as the discussions sparked by the replication crisis show, 'expansion' needn't mean scientific progress as defined by the psychologists themselves. The scientific psychologists' consensus that promised to "finish" psychology – produce comprehensive theory or at least organized knowledge – just could not deliver. I will discuss these fault lines in section 1.3.

1.3 Replication crisis

Thus far, I have elaborated on the institutionalized conventions I am arguing dominated research in psychology in the second part of the 20th century. The recent replication crisis is an event destabilizing those conventions. In an extreme reading it could even be seen as something destructive toward them, especially for psychology's institutionalized practices of using inferential statistics. The 2010s replication debates

[37] Here, I only talk about the genre of the research article. My analysis could easily include the development and importance of literature reviews, as genres that organize larger numbers of empirical articles and serve a crucial role in systematizing the incremental encyclopedism of varied psychological communities (for literature reviews as a type of "scientists' imagined pasts", see Blum, 2017). Not only that this is a separate genre in 20th century scientific psychology, but it is also serviced by specialized overview journals that publish only those sorts of reports, e.g. *Annual Review of Psychology*. Another practice fulfilling the same function are statistical procedures like meta-analyses, and consequently, meta-analytic articles.

are a deep controversy that is still developing and affecting psychology's subdisciplines, as well as other social and biomedical sciences. In this introduction, I will use it to highlight and critically discuss the previously described conventions of scientific psychology.

First and foremost, what is the replication crisis? It is a collection of critical debates that arose in the 2000s and 2010s about the diminishing trust in existing scientific literature in potentially many scientific disciplines, psychology being just one of them. It is important to stress that the discussions, for the most part, are not anti-science – they *are not* formulating a claim that science should be mistrusted *tout court*. Rather, most of the vocal critics in the discussions are reformers. They are practicing scientists who are raising serious issues about methodology, statistics, and epistemology in their research fields.

The standard early articulation of this worry over the trustworthiness of scientific literature came from John Ioannidis, an epidemiologist working at Stanford University. In 2005, Ioannidis published a paper under the title *Why most research findings are false*, in which he argued that "there is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims," (p. 696) and even more so, that this concern can be proven empirically, by investigating how scientists conduct and publish their research.

Ioannidis goes on to argue that as a consequence of small sample sizes, small effect sizes under investigation, preference by journals to publish only positive results, flexibility in research designs and data analysis, perverse incentives pushing scientists to publish more and faster, and the gold standard of p-values as a criterium for publication; whole scientific fields are potentially just collections of biased estimates of true effect sizes. In other words, the phenomena that get to enter the literature are filtered in an extremely biased way that does not actually select for true descriptions of the world. At first, this criticism focused on biomedicine. But other fields were potentially afflicted with the same problem, as Ioannidis and his research group seem to have been aware when founding their Meta Research Innovation Center at Stanford. They proposed tackling the problem by founding a new scientific discipline that deals with it, a sort of 'science of science' that will empirically identify the problems and provide solutions (Ioannidis, Fanelli, Dunne, & Goodman, 2015).[38]

Considering the efficiently institutionalized inferential revolution among scientific psychologists in the 20th century, it did not take long for certain subdisciplines of psychology to become a poster child for untrustworthy science that Ioannidis and the meta-scientists were talking about. Scientific psychology, with its bureaucratized methods and statistics, was ripe for the picking by meta-scientists who almost

[38] The relationship between the developers of "science of science" and the replication crisis reformers seems to be ambivalent at times. For an example, see Fanelli (2018).

exclusively focused on methods and statistics. This was also facilitated by a number of high profile controversies concerning outright fraud and questionable research practices by reputable psychologists. The two with the widest reverberations and public outcry were the 2011 publication of a study by Daryl Bem that claimed to statistically prove precognition in the *Journal of Personality and Social Psychology* and the case of systemic fraud by a Dutch star social psychologist, Diederik Stapel. Since the Bem and Stapel episodes were two events that pushed psychology into the center of the replication crisis, I will shortly discuss both of them.

1.3.1 Foreshadowing the crisis: Precognition and shaky methodological standards

Daryl Bem[39] (b. 1938) is one of the more prominent American social psychologists in recent history. He studied at Reed College, Massachusetts Institute of Technology, and obtained a PhD degree from University of Michigan in 1964. He taught at Carnegie Mellon University, Stanford, Harvard, and Cornell from which he retired and became an emeritus in 2007.[40] In 2011, as an emeritus, he published an article in which he argued that he had statistically proven precognition – the ability to see into the future, or as he operationalized it in his studies, to test whether the participants of his studies had the ability of "detecting a future event" (2011, p. 409).

Bem's experimental studies were pretty ingenious. He involved more than a 1,000 participants across the nine reported experiments. In the experiments, he used the then standard social psychological experimental paradigms. Experimental paradigms are standardized set-ups for experiments that are used by social psychologists to investigate the constructs they are interested in. The only difference Bem introduced was "time-reversing" them. Time-reversing meant that "the individual's responses are obtained before the putatively causal stimulus events occur" (Bem, 2011, abstract). He redesigned four experimental paradigms in this way: Approach and avoidance of negative stimuli, priming, habituation, and facilitation of recall. After collecting the data from such parapsychologically altered experiments, Bem just conducted analyses and drew conclusions in the way that was the standard good practice of his field in 2011 – so standard and so good that it got published in one of the most respected venues for publishing on social psychology.

Immediately after the publication of Bem's precognition paper, psychologists at large realized that a respected professor in social psychology used standard methods in their field to statistically prove that people could see into the future. The publication of the

---

[39] Basic biographical information about Bem is retrieved from his Wikipedia page.

[40] Bem is also one of the co-authors of a highly successful American undergraduate psychology textbook, *Hilgard's Introduction to Psychology*, that will be analyzed in Chapter 3. Bem wrote the chapter on social psychology in the textbooks' later editions.

paper provoked a fierce debate among psychologists. Most were extremely skeptical about any kind of *psi* phenomena receiving scientific corroboration, and yet, all the proxy indicators of scientific robustness were saying that these results should be at least entertained as pointing toward the possibility of precognition being real. Bem was well-trained, worked at some of the most prestigious universities, and the journal was one of the best. Considering this, there must have been something wrong with the study or the scientific standards, because not many of Bem's fellow psychologists were happy to include precognition among the constructs in circulation in their discipline's literature.

Re-analysis, criticism, and failures to replicate Bem's results soon followed (Wagenmakers, Wetzels, Borsboom, and Van der Maas, 2011; Rouder & Morey, 2011; Ritchie, Wiseman, & French, 2012). *Psi* phenomena were swiftly disinfected and removed from mainstream psychology, putting them on the contested margins where they have existed since the 19th century.[41] But the doubt remained. It was articulated well by a commentary published by Etienne LeBel and Kurt Peters in 2011:

> Bem (2011) deserves praise for his commitment to experimental rigor and the clarity with which he reports procedures and analyses, which generally exceed the standards of MRP [modal research practices, which is the authors' shorthand for the accepted methodology empirical psychologists most commonly use in their research] in empirical psychology. That being said, it is precisely because Bem's report is of objectively high quality that it is diagnostic of potential problems with MRP. By using accepted standards for experimental, analytic, and data reporting practices, yet arriving at a fantastic conclusion, Bem has put empirical psychologists in a difficult position: forced to consider either revising beliefs about the fundamental nature of time and causality or revising beliefs about the soundness of MRP. In this commentary, we explore the possibility that deficiencies in MRP can indeed provide an alternative explanation for the publication of Bem's article. (p. 371)

Where Bem failed in persuading most of his fellow psychologists about the existence of *psi* phenomena, he succeeded in casting a shadow of a doubt on their research practices. And in the following years, that doubt would grow into an elaborate movement of criticism and reform. Bem's paper was published online in January of 2011. The debates about it – among psychologists and the wider public – raged throughout the year. But what made the crisis of confidence truly explode was a fantastic case of scientific fraud a-brewing in the Netherlands as 2011 was nearing its end.

---

[41] For the history of the belief in extraordinary psychological phenomena and its relationship to expertise, see Lamont (2012) and Sommer (2012; 2014). An interesting sociological study of scientific replication and parapsychology can be found in Chapter 5 of Harry Collins' book *Changing Order* (1992).

1.3.2 Diederik Stapel's publication factory[42]

In September of 2011, a huge case of scientific misconduct came to light at Tilburg University, a research university in the south of the Netherlands. Their star experimental social psychologist and dean of the School of Social and Behavioral Sciences was suspended after a special commission concluded he committed extensive fraud in many of his publications. The nature of the fraud was fantastic – Stapel manipulated or even completely fabricated datasets used in dozens of published studies. He, up to that point, had been one of the most successful Dutch social psychologists. This is how his research was described on Retraction Watch (2011), the international watchdog for retractions in science, as the story of his fraud broke out:

> His articles, on everything from table manners to infidelity, have been published in both the social science literature and more general titles, including an April 2011 paper in *Science* on discrimination, and he has collaborated with researchers in both Europe and the United States. Twenty eight of his papers have been cited at least 20 times, according to Thomson Scientific's Web of Knowledge, and two have been cited more than 100.

Following his suspension because of confirmed suspicions of data fraud based on the investigation of the special committee at Tilburg University, the original committee cooperated with the University of Amsterdam and University of Groningen, at which Stapel worked during his career, to launch parallel investigations. By the end of 2012, the now three committees in Tilburg, Amsterdam, and Groningen concluded their investigations of the Stapel case with a 104 page report (Levelt, Drenth, & Noort, 2012). The committees, with help from statisticians funded by the three universities, combed through Stapel's publications, looking for fraud. Their conclusions were staggering. They established that 55 publications by Stapel were fraudulent.

None of his PhD students or co-authors were implicated in the fraud, but the fallout caught many who collaborated with Stapel (Levelt, Drenth, & Noort, 2012): "In a formal sense, the people affected are hampered in their careers, such as when extending temporary contracts and applying for grants. [...] In an informal sense, there is an element of stigmatization that may persist long into their further career. For some victims the consequences may be more drastic than they are yet able to foresee" (p. 34).

The scale of the fraud resulted in serious doubts about the scientific standards of social psychology, or even scientific psychology as a whole. As Abma puts it: "The unmasking of Stapel spelled bad news for social psychology. In the public opinion, the discipline was ridiculed by columnists and writers of letters to the editor of newspapers [...],

[42] The subtitle is taken as a direct translation of Ruud Abma's (2013) book-length analysis of what has become known as the Stapel affair. Abma's work has informed most of this subsection.

some even recommending to shut down psychology as a whole. Partly due to the extent of Stapel's fraud, the credibility of social psychology also came under pressure in the academic community" (2013, p. 125).

This was far from an incident limited to Dutch psychology or educated public. The Stapel affair was part of the context in which the Nobel Laureate Daniel Kahneman (2012, September 2012)[43] wrote a scathing open letter directed at priming research in social psychology:

> As all of you know, of course, questions have been raised about the robustness of priming results. The storm of doubts is fed by several sources, including the recent exposure of fraudulent researchers, general concerns with replicability that affect many disciplines, multiple reported failures to replicate salient results in the priming literature, and the growing belief in the existence of a pervasive file drawer problem that undermines two methodological pillars of your field: the preference for conceptual over literal replication and the use of meta-analysis. Objective observers will point out that the problem could well be more severe in your field than in other branches of experimental psychology, because every priming study involves the invention of a new experimental situation.

Kahneman also mentioned publication bias – an old problem that was resurfacing as the trustworthiness of psychological literature was becoming questioned. Already in 1979, Robert Rosenthal had described the file-drawer problem: The issue that only positive results i.e. corroborations got published, while studies with negative results remained in the psychologists' drawers, gathering dust. This practice leads to a systematic bias in the literature, which becomes extremely problematic when psychologists and other scientists start producing quantified estimates of effect sizes based on the published literature.[44]

Since 2012, the crisis of confidence has not subsided. Rather, it spread to other subdisciplines. Researchers started raising questions about standards of research in scientific psychology at large. Stapel's case was shocking, but it could not be only interpreted as an individual's moral failure. It also exposed the lax standards of journals and hinted at structural problems in the norms and practices institutionalized among researchers in psychology. Problems were not just individual breaches of integrity, but widespread systemic malfunctioning of the science system. As the editors of a Special Section on the Replicability in Psychological Science of

---

[43] See Ed Yong's (2012) piece on Kahneman's letter in *Nature News*.

[44] This has become a large topic in the 2010 debates. For elaborations of the publication bias, see Francis (2012), Ferguson and Heene (2012), and Bakker, Van Dijk, and Wicherts (2012). For an influential attempt of tackling the file-drawer problem, see Simonsohn, Nelson, and Simmons (2014) on p-curve analysis.

Perspectives on Psychological Science put it, this was not only a crisis but an opportunity for wider reform (Pashler & Wagenmakers, 2012):

> In the opinion of the editors of this special section, it would be a mistake to try to rely upon any single solution to such a complex problem. Rather, it seems to us that psychological science should be instituting parallel reforms across the whole range of academic practices—from journals and journal reviewing to academic reward structures to research practices within individual labs—and finding out which of these prove effective and which do not. We hope that the articles in this special section will not only be stimulating and pleasurable to read, but that they will also promote much wider discussion and, ultimately, collective actions that we can take to make our science more reliable and more reputable. Having found ourselves in the very unwelcome position of being (to some degree at least) the public face for the replicability problems of science in the early 21st century, psychological science has the opportunity to rise to the occasion and provide leadership in finding better ways to overcome bias and error in science generally.

The criticism and attempts to reform culminated in 2015, with the publication of a massive cooperative replication study by the Open Science Collaboration (2015).

1.3.3 Psychology's 21st century crisis of confidence

In 2015, four years after Bem's article and three years after the Stapel affair, Brian Nosek collaborated with 269 colleagues from around the world to publish what was probably one of the largest collaborative efforts in the history of psychological research. They attempted to replicate the results of a hundred studies (both experimental and correlational) from three journals - *Psychological Science*, *Journal of Personality and Social Psychology*, and *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Open Science Collaboration, 2015), with the goal of an "initial estimate of the reproducibility of psychological science" (p. 943) The widespread interpretation of the results was that "while 97% of the original studies reported significant results only 36% of the replication studies did so" (Fabry & Fischer, 2015). The large study caused discussions and debates that are still ongoing. Just looking at the amount of attention it garnered on the Internet gives us the contours of its impact. As measured by Altmetric[45] in March of 2018, the Open Science Collaboration paper

---

[45] Altmetric is a dedicated service for tracking the impact of scholarly research outputs. Looking at the geographical distribution of tweets mentioning the OSC study gives us an estimate for the current spread of the conventions of scientific psychology described in this thesis. In March, 2018, the top nine countries according to number of tweets mentioning the study were: United States (500), United Kingdom (196), Netherlands (78), Spain (61), Japan (53), Canada (53), Australia (48), Germany (40), and France (28), or, in other words, the contours of the Global North. This should of course not be taken as a definite description of the spread of scientific psychology, just an indicator. Another such representation of the spread of

was mentioned in 152 news outlets, 113 blogs, by 1961 Twitter users, 6 Wikipedia pages, 25 Google+ users, and 8 Redditors.

What followed its publication were endless discussions: About the exact estimate of how many of the included studies actually were replicated (Gilbert, King, Pettigrew & Wilson, 2016; Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016; Inbar, 2016; Johnson, Payne, Wang, Asher & Mandal, 2017; Bench, Rivera, Schlegel, Hicks & Lench, 2017; Hartgerink, Wicherts & Van Assen, 2017; Etz & Vandekerckhove, 2016),[46] whether we can conclude, based on effects failing to be significant in one study, that they are non-existent (Maxwell, Lau, & Howard, 2015; Earp & Trafimow, 2015; Fiedler & Prager, 2018), fierce debate about the difference between exact/direct and conceptual replication (Stroebe & Strack, 2014, Crandall & Sherman, 2016), the role that lack of robust theory plays in failed replications (early discussion by Klein, 2014; and a rebuttal by Trafimow & Earp, 2016), the rehashing of the significance testing and p-value threshold debate (Benjamin et al., 2018; Lakens et al., 2018; the American Statistical Association also issued a policy statement on p-values, see Baker, 2016) and so on. The debate whether there is a replication problem also spread from social psychology to other subdisciplines of psychology and beyond (e.g. Coyne, 2016; Szucs & Ioannidis, 2017; Schweizer & Furley, 2016), engulfing everything from sports psychology to cognitive neuroscience.

Growing around these debates - of which I cited only the end- and starting-points in journal articles and commentaries, while the fiercest discussions happened on social networks and blogs - was a reform movement. Driving it were a few initiatives. The largest one, that acts as an organizational/infrastructural framework and community builder, is the Center for Open Science (COS, https://cos.io/). COS was launched by Brian Nosek and Jeffrey Spies with funding from the Laura and John Arnold Foundation.[47] Most other initiatives in the reform movement are either directly or indirectly connected to COS. One of the most prominent of such initiatives is the call for instituting registered reports in psychology journals (Wagenmakers, Wetzels, Borsboom, Van der Maas & Kievit, 2012; Chambers, 2017).

Registered reports are a new kind of publishing practice, at least for psychology, in which the scientist registers the research and analysis plan with a journal before collecting data. The design is peer reviewed and either accepted or rejected before the data are actually collected, alleviating many of the methodological and statistical issues raised in the replication controversy. This also relates, as one implementation,

---

scientific psychology is the world map of labs participating in the *Psychological Science Accelerator* (Moshontz et al., 2018). The accelerator is another collaborative effort of the reform movement.

[46] This list of references to re-analyses of the data from the OSC 2015 study was taken from the Tilburg Belief Systems Lab blog (2017, November).

[47] For an detailed timeline and list of activities, see: https://cos.io/about/brief-history-cos-2013-2017/

to the wider initiative within the Open Science movement to reform how peer review and journal publishing works in the 21st century (Nosek & Bar-Anan, 2012; Tennant et al., 2017).

The movement also found expression in the founding of the Society for the Improvement of Psychological Science (SIPS; http://improvingpsych.org/), which, as its mission statement says, acts as "a service organization aiming to make psychological science higher quality and more cumulative." Such developments have prompted different ways of largescale collaborations in conducting psychological research, like ManyLabs projects (e.g. Klein et al., 2014) and Psychological Science Accelerator (Moshontz et al., 2018).

1.3.4 The replication crisis as a magnifying glass for investigating scientific psychology

Just this brief and limited excursion in trying to grasp the kind of debates spawned around replicability of findings in psychology in the 2010s already shows how large the discussion is, with many moving parts. Many statistical, methodological, epistemological, and normative issues have cropped up, with a myriad of proposed reforms and solutions to fix them. As Green (2018, February, p.1) put it when reviewing the problems raised in the replication debates:

> Psychology is currently facing a number of quandaries of a methodological or statistical nature. Any one of them would probably be solvable but, in combination, they threaten to undermine the credibility of the discipline as a whole. What is worse, none of them is solvable in isolation because they interlock with each other, reinforce each other, and make it difficult to solve one without simultaneously solving all of them. Solving all of them would require a virtual revolution in the culture of scientific psychology: everyone moving in the same direction – a direction that will make things a little more difficult and complicated for everyone – all at the same time. Needless to say, in a discipline as vast, diverse, and fractious as psychology, that would be a mammoth undertaking and, frankly, I'm not sure that we're up to it.

Green voices a very important concern. Despite the fact that the crisis and reform debates are prolonged and large, they are not affecting most psychologists. An argument could be made that the issues are being discussed at the forefront – the kind of avantgarde of methodologists and iconoclasts clashing with prominent researchers from prestigious universities. That kind of a clash produces the most publicity and noise,[48] lands one publications in *Science* or *Nature,* and interviews in popular magazines. On the other hand, it's an open question how far it trickles down and

[48] This is precisely what happens if one follows the "tone debate." See Morawski, (in preparation) and a popular coverage of the case of Amy Cuddy in the *New York Times* (Dominus, October 18, 2017).

affects the practices of potentially hundreds of thousands of psychologists conducting research worldwide. Scientific psychology is much larger than the small number of participants in the reform debates, and it is still hard to say whether the movement will spread further.

The reason why I included the replication crisis in this introduction – and discuss it at length in Chapter 5 – is that the copious amount of writing produced in these debates can be used to expose and explicate psychologists' epistemological commitments. The methodological and statistical discussions, mixed in with normative talk of the functioning of science as a system, are an analytical searchlight for a historian or a philosopher. Especially for one that wants to describe the institutionalized conventions of scientific psychology in recent history, like I am trying to do. Looking at these debates reveals many otherwise implicit views held by psychologists about the practice and goals of their research and the state of their theories and methods.

Put in context of the inertia of scientific psychology I have described in section 1.2, the replication crisis can be seen as an accumulation of psychologists' ignored oppositions during the 20th century. A kind of rattling of intellectual skeletons in the closet. In that reading of the replication crisis, the late 19th/early 20th century ignored opposition toward the idea that psychological phenomena can be measured, the mid- and late-20th century ignored opposition toward the idea that operationism serves as a sound approach to defining and refining psychological phenomena, and the vehement but sidelined late-20th century ignored opposition to the institutionalized usage of null hypothesis significance testing; all of those oppositions are gaining new voices. Of course, for the most part, not directly – for example, I haven't read a single reform text from the 2010s arguing positively (or in any way, for that matter) for Kant's in principle rejection of quantification of psychological phenomena. Such an example would directly connect the historical ignored oppositions to the contemporary debates, and since I doubt any exist, this is a position I would be very reluctant to argue for. Rather, by saying that those old ignored oppositions are being voiced again I mean that they latch onto the same fault lines in the big and inconsistent project of scientific psychology in the 20th century. Scientific psychology might be widely and strongly institutionalized through its conventions, but these conventions that follow fault lines make for a fragile science.

If I am correct in that view, conceptual work in history and philosophy of science could be very valuable in reframing and pushing the disjointed debates that have been raging for years. Today's avantgarde critics are grasping for a "robust" and "cumulative" science while the silent masses of psychological researchers are waiting for the dust to settle, to pick up the things that can be implemented in their fields and practical conditions of research. The inertia of scientific psychology has expanded for decades mostly uninterrupted, but now, the window for constructive change is left ajar. My argument is that the inertia of scientific psychology is extremely large, and it will take a philosophically and historically informed reform to truly change it. Just increasing statistical and methodological sophistication will not do. This finally leads me to explicating the research questions and the chapter structure of this thesis.

1.4 Outline of the thesis

The thesis consists of four chapters and a conclusion organized in three themes that investigate 20th century scientific psychology and its inertia. The chapters were written as journal articles, thus they represent the developing view on scientific psychology during my research. They are topically organized into themes: Chapter 2 and 3 on textbooks, Chapter 4 on psychology's scientific literature, and Chapter 5 on psychologists' indigenous epistemologies.

Textbooks. Chapter 2 and 3 look at psychology's textbooks, discussing the role of education in the inertia of scientific psychology. In Chapter 2, I describe two communities of scholars who write about the history of psychology's textbooks. The first are psychologists who use textbooks in teaching, and write histories of textbooks as amateurs i.e. scientists who do not have history of psychology as their primary research interest. The second are historians of psychology and historians of science, who in their historical research use textbooks as sources. This chapter serves three purposes: It sensitizes us to the difference in perspective between historians and psychologists, it provides a historiographical overview of what was written about the history of introductory textbooks in psychology, and it implicitly showcases how the conventions of scientific psychology act in directing the way psychologists approach their objects of research. Teachers of psychology writing about history of textbooks inadvertently produce the kind of research that conforms to the conventions of 20th century scientific psychology. Textbook content is yet another construct that needs to be operationalized in order to understand it.

In Chapter 3, I take a single influential introductory textbook to psychology, Hilgard's *Introduction to Psychology*, and analyze the way it introduces psychology as a scientific discipline to students. Since the same textbook was continuously re-published since the 1950s, I consistently traced the discourse that reproduces scientific psychology.

Psychology's scientific literature. Chapter 4 looks at psychology's scientific literature in the period from 1950 to 1999, providing an empirical largescale description of the discipline using its literature as a proxy. In collaboration with the scientometrician Nees Jan van Eck, I data-mined the abstracts and titles of more than half a million research articles published in hundreds of psychology journals. The data mining was conducted in order to construct a series of term maps of psychology's literature. We analyzed the structure of these term maps as a proxy for the discipline of psychology, discussing it in light of Lee Cronbach's influential view that scientific psychology was separated into two methodological traditions – correlational and experimental psychology. I take the stability-through-time of the analyzed literature as one of the manifestations of scientific psychology's inertia.

Indigenous epistemologies. In Chapter 5, I discuss psychologists' indigenous epistemologies, focusing on the epistemological writings of the reformers in the replication crisis debates. I take the analytical concept of indigenous epistemologies and describe it in two historical episodes. The first is the episode it was invented for by Laurence Smith (1986), namely, the interaction between neobehaviorism and logical positivism earlier in the 20th century. The second, which I extend it to, are Abraham Maslow's views on science in the 1960s. The bulk of the chapter is dedicated to analyzing the indigenous epistemology of the reform movement – the kind of notions about scientists as individuals and science as a system psychologists express when discussing the issues of replication and replicability in 2010s.

The thesis as a whole has two aims. The first is to produce a historically informed description of scientific psychology since the 1950s in two ways: By using the discourse of introductory textbook writers (Chapter 2 & 3) and the data-mined language of psychologists in the abstracts and titles of journal articles they published in the period from 1950 to 1999 (Chapter 4). Focusing on the language of psychologists allows me to interpret general trends about the way psychologists construct facts and argue about the content matter of their science. It is also consciously used as a strategy that will identify whether there is a general trend, or as I prefer to call it, the inertia that structures how psychologists write, think, and argue.

The second aim of the thesis is to engage in conceptual analysis of scientific psychology. I do this by first describing the ideas about what are the goals and practices of scientific psychology as argued by the reformers – psychologists involved in the debates about the replication crisis. In this way, in Chapter 5, I retrieve some of the psychologists' basic intuitions about scientific practice, statistics, theory, and scientific progress. I then use my description of the inertia of scientific psychology to criticize the reform movement on epistemological grounds. My aim is to reframe the reform debates as a collection of epistemological and historical questions that are more complex than just the issue of misapplication of methods. In the thesis' conclusion, both the empirical part of the thesis (Chapter 2, 3, 4) and the more critical one (Chapter 5) are rephrased as a call to historians, philosophers, and psychologists for more thorough and collaborative conceptual analysis of scientific psychology.

Chapter 2. Instructional manuals of boundary-work

In the past years, there has been a boom in the history of science pedagogy, and with that, textbook history. Kathryn Olesko (2006, p. 863) unequivocally states that "[t]he historical study of science pedagogy has of late experienced a renaissance, and with that, a revolution in perspective." The articles in the volume of *Science & Pedagogy* for which Olesko wrote the above-cited commentary attest to this change in perspective and focus, bringing together a collection of studies on the history of scientific and technological textbooks on the European periphery. As Bernadette Bensaude-Vincent (2006) argues in the same volume, science teaching in general is interesting for historians "[n]ot only because it is indispensable for training new generations of scientists or because it enriches our view of science as a social and cultural activity, but also because it determines the disciplinary partitions of scientific knowledge." She continues with saying that textbooks in particular "are sorts of archeological traces of former regimes of knowledge" (p. 668). Textbooks are not just repositories of uncontroversial facts used to disseminate them outside of tight-knit scholarly communities— they can also serve as epistemological and institutional catalysts.

Olesko, and to a certain extent Bensaude-Vincent, argue that the once-exciting interpretative frameworks of Thomas Kuhn, Jerome Ravetz, and Michel Foucault have been eclipsed, at least in the case of the history of science pedagogy. Olesko (2006) states: "[U]nderstanding pedagogical experience is something greater than the sum of institutional and intellectual history" (p. 871). The gloomy perspective formed at the intersection of Foucault's disciplining of disciplines in *Discipline and Punish* and *The Order of Things* on one hand, and Kuhn's textbook science and its role in paradigmatic establishment of "normal science" on the other, is too sweeping and too insensitive to the particulars of what was actually happening in science classrooms in the eighteenth, nineteenth, and twentieth centuries. We need not go far to look for an example of the rich context these perspectives tend to stamp out, for example, Andrew Warwick's (2003) excellent study of the role of the pedagogical context and its tools in the rise of mathematical physics at Cambridge.

Science pedagogy, in particular concerning textbooks, is also relevant for history of psychology as a discipline. In this chapter I will focus on the historiography of introductory textbooks in psychology, highlighting the differences in the approaches taken by psychologists with those taken by historian-psychologists. I aim to show that these textbook historiographies are particularly interesting as a space where genres of different communities meet to describe the same things—the textbooks used in introductory university courses in psychology—but the resulting descriptions are quite different. We can see the different agendas psychologists and historian-psychologists project into their narratives, and in turn, the object constituted by those narratives is quite different. The difference results in a divergence in historiographies—two parallel streams of thought talking about introductory textbooks in psychology.

Psychology textbooks are not only interesting because of the divergence in historiographies talking about them, but also because of the particularities of textbooks in psychology versus other natural or human sciences. In psychology, textbooks are not only actual physical artifacts containing the partitions and classifications of previous (potentially eclipsed) knowledge systems, they are also a place for the expert to construct new subjectivities for the student (Morawski, 1992, 1996). This is quite particular to textbooks in the social sciences, and especially so in psychology. The authorial voice of the psychologist attempts to offer a different subjectivity to the student than the one experienced in the everyday: A scientifically constructed one.[49] In effect, the twentieth century introductory textbook in psychology is an attempt made by psychologists to offer what are currently legitimate concepts of human psychology to the student. These new concepts aim to be the furniture of the students' private (and idiosyncratic) psychological reality. The scientific account attempts to supplement, or even in a radical reading supplant, the commonsensical one—the one each of us has access to by the virtue of being human beings. This is especially the case in the textbooks found at the end of the nineteenth century, when psychology as a discipline is establishing itself. This makes psychology's textbooks particularly interesting as sources—of disciplinary partitioned regimes of knowledge that construct subjectivities. This construction of subjectivity also presupposes a particular epistemology used in fact making in psychology's textbooks that is not seen in other disciplinary introductory texts (Smyth, 2001a, 2001b, 2004).

The goals of this chapter are to describe two approaches to the histories of introductory textbooks in psychology—one produced by psychologists writing about textbooks what I call the received view; and the alternative view by psychologists-historians (Vaughn-Blount et al., 2009) and historians. The received view is fashioned from a large number of articles published in the journal *Teaching of Psychology*, and then later synthesized into a historical account in Weiten and Wight's (1992) text; while the alternative view is discussed mostly through the work of Jill Morawski and Mary Smyth. I am primarily interested in the historiography of introductory texts used in undergraduate courses teaching general psychology, but the investigation of these

[49] The textbook's goal of constructing a subjectivity that is at odds with commonsense is precisely what Morawski (1996) aims at: "Psychology's success in undermining commonsense knowledge and in marketing an apparently unsavory model of subjectivity depended on the reader's dissociation from that subjectivity" (p. 153). This notion of waging war on the commonsense understanding of the world signals the difficult position psychology is in as a scientific discipline. If we compare it to W. V. O. Quine's (1951, p. 42) notion of "science [as] an extension of common sense," no wonder that "patrolling the borders of psychology" (Smyth, 2001b) as a scientific discipline takes so much effort and specific strategies. It serves as epistemological legitimization; in other words, psychology is either scientific/counterintuitive or commonsensical/trivial.

books is unavoidably tied to other genres of textbooks—subdiscipline specific ones (e.g., social, developmental, abnormal psychology textbooks), or textbooks in methods and statistics.

In comparison to a renaissance of history of science pedagogy and history of textbooks as a broader trend, however, historians of psychology have still mostly avoided textbooks. There is some historical research using textbooks in psychology; some episodes in the discipline's history are entangled with using textbooks as sources—the case of Ben Harris's investigations (2011, 1979) of Watson's *Little Albert* experiment come to mind. The same goes for Stam, Lubek, and Radtke's (1998) investigation of the particular view on the Milgram's experiments produced and ossified in social psychology textbooks. An even more popular genre of textbook research among psychologists is investigating their biases (Brown & Brown, 1982; Winegard, Winegard, & Deaner, 2014) or even just examining them for general inaccuracy and errors (Steuer & Ham, 2008). This, however, is more contributing to the image of the denigrated faulty textbooks and their flawed accounts of science (Morawski, 1992) than to textbook historiography. For a more historical perspective, Thomas Teo's (2007) study of German psychology textbooks in the beginning of the nineteenth century is a breath of fresh air. Andrew Winston's studies on textbook definitions and redefinitions of psychological experiments should be mentioned here (1988; Winston, 1990; Winston & Blais, 1996; MacMartin & Winston, 2000; Winston, 2004), as well as his study on the changes in presentation of race and heredity in introductory textbooks (Winston, Butzer, & Ferris, 2004).

The author that does stand apart in this canvassing of historical work on psychology's textbooks is the previously mentioned Jill Morawski (1992) with her almost programmatic article in the *American Psychologist* on how textbooks of psychology create subjectivities, and her later full analysis of the rhetoric in textbooks in the beginning of the twentieth century (Morawski, 1996). Coupled with a series of Latourian readings of psychology's textbooks from the second part of the twentieth century by Mary Smyth (2001a, 2001b, 2004), and the already mentioned investigations of the experiment through psychological textbooks by Winston (1988; Winston & Blais, 1996; especially the discourse analysis in MacMartin & Winston, 2000), Morawski's approach offers a solid basis for a historiography of psychology's textbooks that fits nicely in the current discourse espoused by Olesko and Bensaude Vincent. Maybe it is not a boom evidenced in other disciplinary histories, but it is a definite presence that must be mentioned.

However, the received historical view of psychology's textbooks is not to be identified with the work of Morawski, Smyth, Winston, Stam, or Teo. As the body of this chapter will show, historiography of introductory textbooks in psychology is quite peculiar. I call it the psychologists' received view of textbooks and attempt to describe it in detail. For the most part, the received view does not delve in the disciplinary partitions found in textbooks. Disciplinary formation and organization of knowledge are not the research objects of the received view. Instead, the type of historical research gathered around the received view on textbooks is precisely an extension of these intradisciplinary negotiations. In simpler terms, the psychologists' methodologies and ways of thinking have expanded from humans and rats to include textbooks. Psychologists have fashioned a received (standard) view of their textbooks that does not have much to do with the work of historians of science. In this chapter, I aim to describe this received view through the example of Weiten and Wight's (1992) large chapter on the history of American textbooks, and a large sample of articles on textbooks published in the journal *Teaching of Psychology*, on which Weiten and Wight build their chapter. Then, this standard view on textbooks will be juxtaposed to that of Morawski and Smyth, and as a conclusion, an integrated approach to the historiography of textbooks in psychology will be suggested.

This integrated view will not only suggest the physical textbooks as a historian's object of research in the case of psychology, but the extended literature about textbooks written by psychologists. Disciplinary negotiations and boundary-work which is interesting for a historian of psychology does not only happen on the pages of introductory texts aimed at freshmen, but also in the wider literature of commentary and research on textbooks done by psychologists. Put like that, the whole genre of writing about introductory textbooks provides an entry point for exploring psychologists' unwritten methodological horizons in general. The way they approach textbooks is also indicative of the way psychologists approach other subjects and research questions, at least in the second part of the twentieth century.

2.1 The received view of psychological textbooks: From whose vantage point?

The received view of textbook history in psychology is produced by a different community of scholars who have nothing to do with tracing disciplinary formation in a way historians like Bensaude-Vincent or Olesko aim at, or identifying the construction of subjectivities like Morawski does. This is not to say that there is an essential tension between this psychologists' received view and the historians' discourse on textbooks sketched thus far—it is more like a chasm of silence. The received view was produced as an unexpected amalgamation, a sort of a big overview that tried to create a historical discourse out of the collective research of a number of scholars publishing on textbooks in the journal *Teaching of Psychology*. The view took actual shape as Wayne Weiten and Randall D. Wight's (1992) chapter on textbooks in the APA volume *Teaching of Psychology in America: A history*.

Other than providing the most detailed chronology and bibliography of English language psychology textbooks ever published, it also provides a particular way of understanding what a textbook as an object of research is, and what it can offer us. The textbook and its history, as described by Weiten and Wight, stand in stark contrast to the discourse sketched from the work of Olesko, Bensaude-Vincent, Morawski, and Smyth. Incidentally, the one time (out of the two) when Weiten and Wight's chapter cites Jill Morawski's work is to corroborate the authors' claim that fellow academics look at textbooks with "suspicion and scorn" (Weiten & Wight, 1992, p. 487). Her substantive contribution to textbook historiography is ignored.

Calling the appearance of the received view of psychological textbooks an unexpected amalgamation without an explanation might be misleading. The production of this view was not unexpected in how it was written or researched—its meticulous and large bibliography attest to that. The unexpectedness was in its entrance to the scene of history of psychology, where it appeared in a chapter in an edited volume on the history of teaching of psychology, tying together a literature that was probably never imagined to produce historically relevant knowledge. The chapter was the crest of hundreds of articles on textbooks published in *Teaching of Psychology*, which in the way they are written, in the points they argue, and in the audience they address do not have much to do with the typical questions historians ask or the answers they hope to get.

To avoid ambiguity—and scholasticism in dispelling it—this point has to be carefully presented because it relates to the crux of the argument in this article. It boils down to the juxtaposition of two histories—one produced for historians and the other produced for psychologists-turned-historians.[50] This is not to say that one of those is "good" history and the other "bad," just that they serve different goals and different audiences. As Bert Theunissen (2001) argues from the perspective of a historian of science—there is no reason why scientists' histories of their disciplines should be "bad" histories by default "[b]ut in practice, scientists' histories do tend to differ quite substantially from the kind of history written by professional historians of science" (p. 148).

Theunissen draws the distinction between scientists and historians, and the aims they have in writing histories. In the case of textbooks, this distinction is complicated by the fact that the group juxtaposed to the historians is not just scientists, but science teachers. Teachers of psychology have a particular goal when they use history—the tension between the history for history's sake on one end and history used to teach on the other leads to Stephen Brush's (1974) complex point in the article *Should the History of Science Be Rated X?* that problematizes the practice of exposing students to narratives that "threaten" scientific objectivity. I argue that this tension is not only manifested in the history found *in* the textbooks, but also in the history *of* those

[50] The line could be drawn by demarcating who produced the histories, but making the distinction between the aimed audiences of the histories is much more fruitful in this case, because both Jill Morawski and Mary Smyth are psychologists by training, but they argue quite different points than the psychologists publishing in *Teaching of Psychology*. The difference stems precisely out of the different questions they ask. These questions are provided with answers that would interest a specific audience, in this case, of historians of science. The audience makes all the difference for the type of argument one tries to construct and make believable, as will be argued in this article. So instead of saying "psychologists' history of textbooks" and "historians' history of textbooks," it would be better to say the history of textbooks for psychologists, and the history of textbooks for historians, respectively.

textbooks. The ground for criticism is not only the familiar Kuhnian point about textbook histories—the grand simplifications found in textbooks instead of histories—but also the contextualization of the sequence of these grand simplifications. In other words, textbook histories have a history.

Teachers of psychology are the most common users, but also often producers, of the said textbooks. It follows then that they have a goal of carving a niche for the history of textbooks. A niche that should be defended against the scientists' scorn toward their inaccuracies and foibles but also against the historians' approach that has the potential of breaking down the discourse in the textbook in order to look for traces of disciplinary formation, contingencies outside of the discipline, or negotiations of objectivity. Negotiations of objectivity, as Brush warns us, are especially perilous when they find their way into the classroom.

Weiten and Wight's goal—in what I call here the received view of textbooks—is to describe how we ended up with modern textbooks, or in their words, the development of these "portraits of a discipline." In comparison, potentially both the historians' and the scientists' view on textbooks have a deflating note, looking down on them from the pedestals of historical contingency (historians of science) or uncompromising objectivity (practicing scientists). The carving of this niche in the face of scientists' derision is evident from the way Weiten and Wight (1992) conclude the chapter: "[I]t seems shortsighted to evaluate introductory textbooks by the canons of scholarship applied to journal articles" (p. 488).

Opposing it to the historian's perspective is more complicated than just reading it out of their conclusions. After all, Weiten and Wight's chapter styles itself as a historical account. To this end, their historiographical approach will be described in detail, with an analysis of the constituency it represents. Calling it a constituency is a conscious choice, trying to avoid terms that might lead us to over-interpret, terms such as *Denkkollektiv* (Fleck, 1979), invisible college or even a research community (Crane, 1972), or Kuhn's own members of a community gathered around a paradigm (1962/2012). Instead, minimal interpretation will be allowed as to what kind of a community is formed by the authors gathered around Weiten and Wight's chapter on textbooks. The emphasis is on their perspective on textbooks, the received view, as it is fashioned from the interaction of the producers' know-how and the audience's expectations.

2.1.1 Weiten and Wight's portrait of a discipline gleamed from textbooks

The 1992 volume on the history of teaching of psychology in America was imagined as a celebratory, but crucial, contribution to the growing scholarship on the history of psychology. As Charles Spielberger (1992), the APA president at the time, aptly put it in the foreword of the volume: "Although history of psychology is well documented in numerous books and articles, relatively little attention has been directed to examining the history of *teaching* of psychology" (p. xvii). The volume was aimed to fill this lacuna in the historiography of psychology, and provide support and incentive for future research. Indeed, the editors themselves call it a baseline: "Although not meant to be the final word, this collection is reasonably exhaustive and thought provoking, raising questions that provide an impetus for further analyses in the field and a baseline from which to trace developments into the 21st century" (Puente, Matthews, & Brewer, 1992, p. 7). They go on to conclude that this will not only add to history, but also to better teaching of psychology in America. Weiten and Wight's chapter, then, is aimed to fulfill this goal in covering the role of textbooks in the history of teaching of psychology. The tone of the volume and its institutional endorsement is what immediately rings of a received view—it announces that the chapters contained within spell out what is what in the history of teaching of psychology.

As Weiten and Wight put it themselves (1992): "The scholarly literature on introductory texts remains sparse, and the few articles available typically focus on one text or author. We hope our chapter will help to fill this void in psychology's intellectual history" (p. 454). Their chapter consists of two large sections—the first one providing a sophisticated and detailed chronology of more than a century of psychology's textbooks, with a basic periodization and extensive bibliography, and the second giving a quantitative analysis of the textbooks in this chronology. The chronological overview is quite a valuable contribution to the historiography of textbooks in psychology and I will shortly examine it in more detail first.

2.1.2 The definite textbook chronology

The chronology allows for a periodization of textbooks in line with the development of American psychology as a discipline—from the textbooks used in courses of moral philosophy in the last decades of the nineteenth century, to the transition to the "new psychology," the period of conflicting schools and theoretical eclecticism, and finally to the rise of the student-oriented texts in the 1930s–1940s and the encyclopedic texts that were the prerequisite for the homogenization and standardization of the modern textbooks since the 1970s. As an overview, this periodization serves its purpose in providing a factual timeline of publications to be analyzed. However, despite its dearth, the periodization is not without its problems.

The first question that a periodization and a classification should engage with is demarcation—what makes a psychology textbook? Is it an introduction to a self-standing academic discipline, or to a range of topics investigated by particular scholars? In Weiten and Wight's chapter these questions are not addressed. Was a book used in the classroom taught by American mental philosophers or the teachers of Scottish "commonsense" philosophy (Fuchs, 2000) an introductory textbook to psychology, or did it take the new psychology imported from Germany and the laboratories proliferating with that migration to constitute the first textbooks? When did the shift from an introduction of psychological topics to an introduction to a (empirical!) discipline happen? These are precisely the kind of questions that should anchor the overarching chronology of psychology's textbooks, but they are just skirted over by Weiten and Wight (1992) with the sentence: "Textbooks intended to introduce college students to the field of psychology emerged gradually out of the work in moral philosophy during the 19th century" (p. 455). What can be read out of this is an essentialist conception of psychology as a discipline.[51]

The particular kind of history of textbooks created by Weiten and Wight comes in full force in the second section of their chapter, which is the analysis of the meticulous chronology they constructed. To avoid being pedantic, and also to alleviate the potential criticism that my interpretation of the standard view was built on a single text written by two psychologists; the following section of this article will try to explore Weiten and Wight's received view through the connected network of articles produced by their audience. There is a small community of scholars expanding out of the *Portrait of a Discipline* into a citation network. This citation network creates an academic ecology in which the received view flourishes, and for which it was crafted in the first place.

2.1.3 How to identify an audience of a chapter in a book?

The audience was identified in a number of ways—it was actually triangulated through three separate approaches: Looking at the reference list of the chapter, conducting a search in the Web of Science database, and identifying the publisher of the journal. All three things pointed, more or less, to a single journal: *Teaching of Psychology*.

The APA Division 2—the division on teaching of psychology that sponsored and organized the writing of the volume containing Weiten and Wight's chapter publishes a journal. The history of the said journal is also a subject of interest for the volume editors, so a full chapter is dedicated to its founding and historical development (Daniel, 1992). The division shares the name with their journal—*Teaching of Psychology*. Not surprisingly, looking at the reference list of Weiten and Wight's chapter, that journal plays a prominent role there too—at a cursory count, publications in *Teaching of Psychology* are cited about 30 times. This is of course no definite proof that the authors of those publications are the intended audience of the chapter, but it does imply a certain community of scholars writing about the same topic.

Going to Web of Science and searching for the topic "psychology textbooks" one journal came under spotlight—by now, no surprise there, it was *Teaching of Psychology*. The journal ranked first considering the number of relevant publications, containing 145 articles on psychology textbooks. The next in line *was Psychological Reports* with just 27.

By this point, the idea that a constituency of the standard view could be traced through the publications on textbooks in *Teaching of Psychology* took form. Instead

[51] For a nonessentialist historical approach, see Thomas Teo's (2007) treatment of German introductory psychology texts in nineteenth century Germany.

of just using the 145 articles pertaining to textbooks in *Teaching of Psychology* that were found in the first search of Web of Science, the search was redid on that particular journal. My reasoning in this was that maybe some of the publications were missed, and will have been located with a finer-grained search of that particular journal. In total this identified 3,174 entries as the whole corpus of the journal. When the full corpus was refined to only those including the word "textbook" in their keywords, titles, and abstracts, we arrive at 184 publications. Then, the reference lists of those 184 articles were analyzed using *CitNetExplorer*, a program developed by Nees Jan van Eck and Ludo Waltman (2014a) specifically for citation network analysis.

2.1.4 A look at the received view through citations

CitNetExplorer directly builds on Garfield's notion of *algorithmic historiography* which is an approach that tries to implement bibliometric tools of citation analysis in historiographical studies; and Garfield developed a program for such analyses. As Garfield, Pudovkin, and Istomin put it in the description of the said software: "[*HistCite™*] facilitates the understanding of paradigms by enabling the scholar to identify the significant works on a given topic . . . it provides a graphic, genealogic presentation of citational links between them" (2003, p. 400). *CitNetExplorer* is a more sophisticated program offering similar functionality—it is basically used for citation analysis of literature, for example, "for studying the development of a research field over time, delineating research areas, studying the publication oeuvre of a researcher, [and] literature reviewing" (Van Eck & Waltman, 2014b). In our case, it was used to analyze the literature about psychological textbooks in *Teaching of Psychology*.

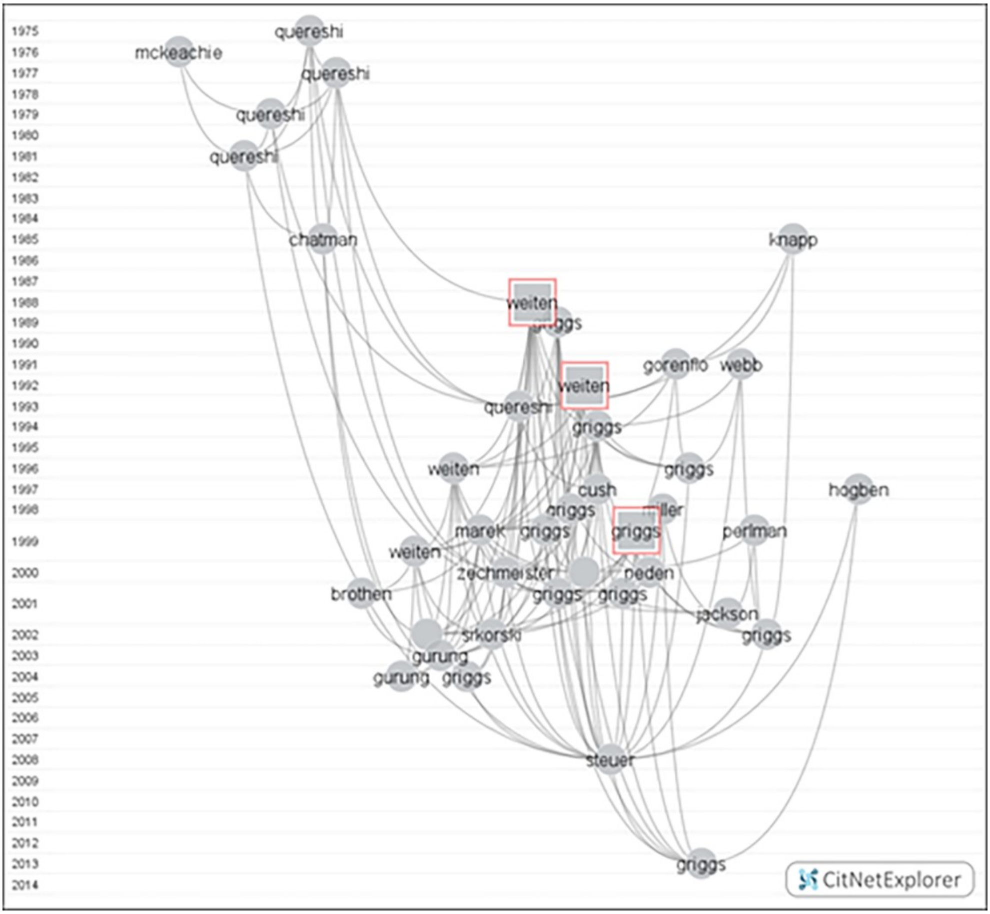


Figure 2.1. Teaching of Psychology citation network

CitNetExplorer extracted references from the reference lists of the 184 articles in *Teaching of Psychology* that were identified to pertain to textbooks, and mapped the references to the articles in *Teaching of Psychology*, and to other publications if they were cited 10 or more times by the said 184 articles. The product of this mapping can be seen in *Figure 2.1*, for clarity only including the 40 articles with the highest citation scores. This represents the framework in which authors publishing in *Teaching of Psychology* write and think when they discuss textbooks. Keep in mind; it is the network of the references of those 184 articles, not of the articles themselves!

The vertical axis represents the publication years. The closer the publications are horizontally, the bigger is their citation relation. Every circle is a single publication, marked with the name of the first author, while the curved lines represent citation relations between the publications. Citations point in the upward direction, meaning that the cited publication is always located above the citing publication. We can identify the central publications in the network according to internal citation scores the number of times a publication was cited within the network. The three most cited publications are marked as squares instead of circles, and those are: Weiten and Wight's *Portraits of a Discipline* (1992; the lower square Weiten in the figure), Weiten's *Objective Features of Introductory Psychology Textbooks as Related to Professors' Impressions* (1988; the upper square Weiten in the figure), and Griggs et al.'s *Introductory Psychology Textbooks: An Objective Analysis and Update* (1999; the only square Griggs in the figure). The full list of publications in the citation network can be found in the Appendix A, with the first 40 publications from the appended list represented in *Figure 2.1*.

All the articles in the Appendix A were inspected to ascertain how many of them are actually about introductory psychology textbooks, versus psychology textbooks in general. This was done by examining the titles, and in unclear cases, abstracts and the articles themselves. Out of the 188 articles in the citation network, 115 deal exclusively or in part with introductory psychology textbooks and courses—others deal with subdiscipline-specific textbooks or other topics relevant to teaching of psychology. It is safe to conclude that the bulk of the focus on textbooks in psychology, as far as *Teaching of Psychology* goes, is oriented toward introductory textbooks.

The interesting thing about this citation network is that it represents the articles and publications on textbooks cited by the authors publishing in *Teaching of Psychology*. So it is not only supposed to represent the articles they publish in that particular journal, but all the articles published across a number of journals that might cover psychology textbook research (if they were cited often enough). Well, in principle that is true, but it actually does not cover anything else but *Teaching of Psychology*. The network is self-contained.

The closed loop is obvious when one looks at where do the publications in the network come from: All except of one are published in the journal *Teaching of Psychology*. The one that was not published in that journal is precisely the Weiten and Wight's chapter analyzed before, sitting there in the middle of the citational mentalscape of the authors publishing about psychological textbooks in *Teaching of Psychology*.

2.1.5 Description of a constituency – Quereshi in the 1970s and the 1980s

Weiten and Wight sampled textbooks from 1890 to 1990 for their chapter, and compared them on a set of quantified variables—text variables, measures quantifying topical coverage, topical organization, book size, illustrations, citations, and references. This is treating the textbook as an object of very familiar quantitative methodologies employed by psychologists in other research areas of psychology. The psychologists' toolbox was applied to an object of research that was made of ink and paper, instead of the objects being—like in that old and weary joke—freshmen and rats.

In our displayed citation network, the oldest consistent example of an approach similar to Weiten and Wight's can be found in the articles published by M. Y.

Quereshi. The content analysis approach pursued by him opened up the field of introductory textbooks to psychologists applying their usual methods of research. Quereshi and his various associates published four articles from 1975 to 1981 on content analysis and its various applications to textbooks (Quereshi & Zulli, 1975; Quereshi & Sackett 1977; Quereshi & Buchkoski, 1979; Quereshi, 1981) that were caught into our citation network as highly cited.

In their first content analysis (Quereshi & Zulli, 1975), the authors stress how important it is to learn about the content of introductory texts, considering they are the biggest influence on teaching of psychology alongside the instructor. Stressing the important role of textbooks seems to be a crucial element of the received view. For the psychologists endorsing it, research on textbooks offers insight into the role and function of classroom education both for the future profession and science. One of the common arguments to support this assertion by the community gathered around the received view are the enormous sales figures, which by the force of their sheer volume assert their importance. For example, Steuer and Ham (2008, p. 160): "Informed estimates suggest that annual domestic expenditures on all psychology textbooks reach $160–200 million (S. Scarrazzo, personal communication, July 11, 2005). Such figures imply that introductory textbooks in combination with those used in upper level undergraduate courses constitute a substantial—and potentially massive—part of the education experience of most psychology students." The other support for the importance of textbooks Steuer and Ham mobilize is Kihlstrom's (2010) survey of "fondly remembered textbooks" on three professional psychology listservs.

Are these compelling arguments for the suggestion that introductory textbooks have a relevant impact on future psychologists, or is this kind of impact reserved for the more advanced handbooks? It is difficult to tell. Winston (1990) questions it when comparing the impact of Woodworth's famous Columbia Bible on experimental psychology to his more popular (by sales) introductory texts for lower level intro courses. As Winston (p. 394) puts it: "Despite these enormous sales [of Woodworth's introductory level *Psychology*], it is difficult to gauge the influence of an introductory psychology course and an introductory text on future psychologists. In contrast, it is in the experimental psychology course that students learn how to do research." Considering the above, the shakily supported (almost taken-for-granted) belief that basic introductory texts do exert influence seems to be an important component of the received view.

Coming back to Quereshi and his associates, we see that he tries to expand on previous studies on comparative readability and human interest by pursuing "an objective and systematic analysis of their [the textbooks] salient terms" (Quereshi & Zulli, 1975, p. 60). The authors' interest is of the practical kind—they aim to provide insight about the content of textbooks that will be used by teachers in introductory courses at American universities and colleges. It is a given, then, that systematic content analysis produced by coding subject indices of 25 textbooks and conducting factor analysis on them, what Quereshi in effect did, will provide useful knowledge about textbooks. The function of the textbook is pedagogical, and Quereshi and Zulli are helping with the role textbooks take in teaching by providing more knowledge about them. That is the logic behind their research. What is the actual execution?

Using factor analysis, they cluster the 25 textbooks around 10 factors, based on the 2,648 terms that appeared in the subject index of more than one textbook. The main part of the article is an extensive explanation of the naming of the 10 factors they have obtained, identifying the similarities between the textbooks gathered around a particular factor. The result of this naming procedure is a taxonomy that topically groups the textbooks.

Even though Quereshi's factor analysis gives us a system of statistically derived categories for sorting textbooks, the procedure of naming the latent variable still involves a close-reading examination of indices, tables of contents, and main bodies of the books to explain the organizing principle derived by factor analysis. When one takes into account the sophistication of this analysis in 1975, involving punch cards and IBM computers—the authors went through a lot of effort to gain an organizing principle that would have arisen by mere canvassing of the books—did we gain anything substantive from this conclusion? Some textbooks cover experimental psychology, others quantitative methods, and yet others focus on genetics. If we reason like this with factor analysis, would we not reason the same by just skimming the textbooks? We probably would have. In that regard, the article fulfills the function of structuring the way a teacher would choose an appropriate textbook for her class. Most importantly, it structures the choice *quantitatively*.

Introductory texts are just aggregators of information—some are focused on one kind of information, some on the other; some are more readable than others, but they all fulfill the same function. The implicit conclusion is also about the method, not about the textbooks—factor analysis is a suitable method to explore this pedagogical function of textbooks, by looking at their content.

The 1981 paper *Analytic Procedures for Selecting a General Psychology Textbook* represents the culmination of Quereshi's approach integrating findings and conceptions of the introductory textbook from his older articles in the cluster. It is also one of the blueprints for the received view of textbook research. The article presents analytical selection strategies the help instructors choose textbooks. The selection strategies are based on quantified variables the author connects with the textbooks in his previous publications—various measures of readability, human interest, the number of pages of text, pages of index, etc. As Quereshi (1981) himself describes the purposes he wishes to achieve: "The present study was done to carry out further analyses, as described below, in order to attain the major objective of devising analytic procedures for textbook selection" (p. 143). It is not only about devising a selection procedure, but about devising an analytic one, which for all intents and purposes means quantified.

The above-described cluster of articles by Quereshi and associates provides a glimpse into the beginnings of consistent application of psychological quantitative methodology to exploring introductory textbooks. The list of articles on textbooks is far from exhaustive (e.g., Gillen, 1973; Harari & Jacobson, 1984), and far from fully contained by the citation network presented in *Figure 2.1*. I do not claim Quereshi and his publications provide a direct model for all the other applications and papers published on the topic. They do provide a case study, and by exploring them in detail we immerse ourselves in the view psychologists consistently endorse and develop when thinking and writing about textbooks. By critically reading the articles and trying to ascertain the goal the authors tried to achieve we are taking a glance at how psychologists went about solving practical problems, like those of textbook selection. What later arose out of the solutions to these practical problems, specifically in Weiten and Wight's chapter, was a history of psychological textbooks.

The same story as the one derived from Quereshi's content analyses could be told out of most articles in the citation network. For example, Weiten's (1988, p. 10) article trying to " [...] gather normative and comparative data on introductory texts and to explore how professors' impressions of these texts may be shaped," or Griggs's report on the change in percentages of certain topical coverages from the 1980s to the 2000s (2014c). It all points to a textbook as a quantifiably dissected research object; a research object that is thriving in *Teaching of Psychology*.

2.2 Alternative to the received view

So far I framed the received view of textbooks as objects of research defined by the "usual psychologists' methodology" and conceptual horizon; textbooks as classroom specific aggregators of information mostly without a wider reflexive content relating to a particular discourse, disciplinary culture, and disciplinary identity. They are pedagogical, they are quantified, and they need to be optimized. This view of textbooks is standardized by a constituency of scholars that exists at least since the 1970s, the view has a journal outlet for publications, and is the baseline for a codified and described history in an institutionally endorsed tome. The historical part was most likely unintended—I do not think Quereshi or most of the other authors in the citation network thought of their research as writing a history of textbooks. They were looking for objective characteristics of textbooks. As psychologists are often wont they looked for objective characteristics of their object of research but what they found was a history (Gergen, 1973). What Weiten and Wight did was explicitly recasting this methodological approach as history.

Calling something usual, like I did with psychologists' methodology, calls for more explanation. The usual psychologists' methodology mentioned here is what Kurt Danziger (1996) calls arithmetization—psychologists turning their objects of scientific interest into quantities represented by variables, "information in a form that lends itself to counting" (p. 30); and by doing so their methodological toolkit becomes an appropriate way to talk about textbooks. The discourse of these techniques then defines the object of research—the discourse of the research practice, its variables and percentages and factorially derived categorical structures; *becomes* what a textbook is. The standard view performatively constitutes the historiographical objects of research, in other words, Danziger's arithmetization totalizes the textbook as an object with particular characteristics.[52] The arithmetization happened already with Quereshi, making textbooks suitable research interests for psychologists, and then twenty years later it was refashioned into a history of textbooks by Weiten and Wight.

Considering the above, an alternative performativity—authors outside of the constituency sketched around the standard view—is sought to allow for an all (or better said, more) encompassing historiographical object, a textbook situated in a time, in a place, and in a community of users and producers; a textbook that tells interesting stories not only if we want to answer questions about the percentage of particular topical coverage through the decades (for example), but also if we want to ask reflective and substantive questions about psychology as a discipline and as a science. Basically, a textbook fit for a fundamentally different audience than the one imagined by the received view. In the following part of the article, the alternative will be described through the work of Morawski and Smyth, identifying its four most salient differences when compared to the received one.

2.2.1 Looking at textbooks as instructional manuals of boundary-work

A detour from the standard view has to start with discursively locating the textbook not only for its quantitatively represented content and the optimization of that content for a classroom, but for it as an object operating in multiple contexts at any given time. Jill Morawski (1992) approaches textbooks with this seemingly nonpedagogical question when she is trying to "ascertain scientists' representation of their work to nonscientists" (p. 161). Not only is a textbook a pedagogical tool in the commonsensical understanding as an assemblage of content that needs to be presented and then learned, it is also an object serving as grounds for exchange between two distinct groups—experts and laymen. It is an instructional manual of boundary-work (Gieryn, 1999) for psychological knowledge. Textbooks are "textual artifacts," as Morawski calls them, which tell us something about the actors' own understanding of psychology, with quite an evaluative judgement included in that from the actors' side. The evaluative judgment being as follows: Because certain content is included in a textbook, it must be a part of a certain canon. It acts as a border between the laymen and the expert, trying to transfer knowledge from one side to the other: From the expert to the laymen, that is, from the professor to the student.

[52] The argument that we sometimes constitute an object of research with the way we research it bears resemblance to Gerd Gigerenzer's (1991) tools-to-theory heuristic. The case Gigerenzer makes is that some cognitive models of the human mind are based on inferential methods. In our case, the example is much less lofty—the descriptive statistics of topical coverage and factor analytically derived categorizations are what the textbook becomes, at least as a historical object.

At the same time, calling it an instructional manual of boundary-work, I mean it actually *does* boundary work on what psychology is through how it is describing the discipline. The negotiations of borders in discipline-making are frozen in textual form.

It is not far-fetched to say that practical examples of textbooks as conceptual workhorses in boundary-work abound—here is an example from the 4th edition of Hilgard and Atkinson's *Introduction to Psychology* (1967, p. 7): "The emphasis in this book will reflect the general orientation of American psychology today, which seeks to place the study of behavior in the context of natural science, recognizing man's affiliation with other biological organisms, and hence tracing continuities between man and other animals. Such an orientation need not deny man's uniqueness where such uniqueness is demonstrable." This quote is taken from the introduction of the book, where it is part of setting the stage—explaining the position and role of psychology in comparison to other disciplines. What this shows is that in psychological textbooks the boundary-making becomes quite explicit, for example, in the introduction of Clifford Morgan's (1956, p. 5) textbook:

> As science increases in pace and widens its grasp, it becomes harder to find the borders of demarcation among the behavioral sciences. Actually, there are no boundary lines, only no man's lands of unexplored territory or overlapping domains in which scientists of different labels work side by side. In the general area of behavioral science, psychology is a kind of meeting ground for the natural sciences such as physics, biology, and physiology, and the social sciences such as sociology, economics and political science.

Reading textbooks as boundary-objects has its merits, but it is far from exhausting their potential as objects of historical studies. Looking for constructed subjectivities these boundary-objects serve to propagate, like Morawski does (1996), sets the bar even higher. One can ask here: Is psychology a science of subjectivities, or of objective laws of human behavior and mental processes? Asking this question is precisely what a thought-provoking historical study of textbooks would motivate, and potentially answer.

2.2.2 Identifying and describing the experts' construal of subjectivity

Looking at psychological textbooks in particular, Morawski sees them as settings for introducing constructed subjectivities—various conceptualizations of the human mind and human experience that are generated by psychologists, but might seem quite artificial to the people reading them. The psychologists' account of the mental and behavioral are not necessarily in lockstep with what we experience every day—this needs mediation, and textbooks serve that goal. Psychologists' constructions of human experience need hard work to expand into the networks of psychological makeups of real people—the personalities, perceptions, intelligences, attitudes, and other concepts psychologists develop do not suddenly appear in people's minds as self-referential terms to describe our own personal psychological experience, there is hard work involved in this conceptual expansion (see Latour, 1999a; Latour, 2005). Hard work done in part by textbooks. As Morawski (1996, p. 146) puts it:

> In advocating a world that takes subjectivity as an object with characteristics not unlike the 'natural' objects of other sciences, and simultaneously claiming superior knowledge of subjectivity, textbook writers had to address and engage the very subjects whose own subjective experiences were to be radically reinterpreted by the science. Textbook authors, then, faced the apparent paradox of denying certain subjectivities while attempting to enlist those very subjectivities in the project of scientific psychology.

This construction of subjectivities is conducted in a particular way—a way of fact-making, basically of persuasively generating knowledge. The rhetoric of the textbook encompasses various strategies of fact making that are quite particular for textbooks in psychology, and we will take a look at them through the work of Mary Smyth.

2.2.3 Exploring the role of textbooks in the construction of facts

Mary Smyth's research on textbooks expands the post-Kuhnian perspective; for her "[t]he textbook, in removing context, is not distorting the history of science, but is actually part of it, and deserves to be studied as part of the continuing process of construction and reconstruction in science (Hacking, 1999)" (2001a, pp. 609–610). She points to an interesting tension— part and parcel to scientific research is decontextualizing and simplifying the object under study. This does not make scientists and the research they conduct unsuitable for historical research; on the contrary, it makes it particularly attractive. Why would such a thing make textbooks unsuitable for historians then? She includes textbooks as one of the building blocks of psychology as a discipline through their role in fact-making, and tries to investigate the thorny path psychological facts travel in their last stage—not at their birth in the laboratory and their trip to the journals like Latour and Woolgar (1986) did—but on their trip from the journal to a textbook.

Smyth is on the hunt for autonomous fact statements in psychology textbooks, with the hopes of answering questions such as: Do the facts produced by psychologists follow the same trajectories as those described by Latour and Woolgar in the case of biology? Are the textual ecologies of those facts comparable between biology and psychology? In effect, is Latour's conception of facts applicable to a different science than biology, in particular, psychology?

In her article *Fact Making in Psychology: The Voice of the Introductory Textbook*, Smyth analyzes psychology textbook chapters on two topics—memory and social interaction (2001a). The conclusion of her analysis is quite provocative: Textbooks in psychology do not function as receptacles of facts like they do in biology or other disciplines. The representation of facts in psychology is different, and this is a consequence of a different fact-making process. Psychological knowledge is presented with the evidence of its making. Unlike biology, where what signals knowledge is the obfuscation of the history of its making, psychologists employ the opposite strategy. Valid knowledge is designated by qualifications. As Smyth puts it: “Psychological evidence carries its knowledge with it,” (p. 628) and this can be extended even further to say that psychological knowledge qualifies as knowledge precisely because of the history of its making.

This is not done to legitimize the substantive fact (e.g., a particular model of memory) but to legitimize the way it was constructed—the methodology that made it possible for that substantive fact to come into being. The autonomous knowledge presented is not about substantive psychological phenomena, but about the ways of reaching and inferring these phenomena. Circling back a bit, this is similar to Quereshi and his textbooks—he is not saying that much about textbooks themselves, but more about the ways one can conceptualize them in a quantitative fashion.

Extending this argument to the extreme, psychological knowledge is then the method employed by psychologist whereas the phenomena, theories, models, and psychological constructs are just epiphenomena to mask the true epistemological claim. The qualifier (the modality in Latour’s words) takes center stage, not in making an ontological claim about human psychology, but to make a strong epistemological claim that psychologists have methods to uncover relevant knowledge about the psyche at their disposal. The textbook is making a claim about the road to knowledge; the actual knowledge at the end of it is of lesser importance.

2.2.4 Problematizing the function of textbooks as vindicators of psychology as a science

In looking at how psychologists as textbooks authors go about presenting their science, Smyth tries to learn something about psychology as a discipline. She puts it in the context of Gigerenzer’s surrogates for theories (1998a) and Bazerman’s (1988) “behaviorist shift” in how psychologists write and conduct their science. The practice of textbooks presenting results and studies with full references precisely fulfills the goal of the textbook author; the goal being the provision of epistemological legitimacy to the discipline where otherwise it would be questioned. The question whether this legitimacy building strategy succeeds is something to be found outside of the textbooks.

To put it differently, the surrogates for theories as Gigerenzer calls them, in our analysis based on textbooks, is the discourse of an empirical psychological science. It teaches epistemologically, it disciplines and orders: “Always present evidence—this is the message about practice that the new psychologist absorbs from these textbooks. The paradigm is one of doing, not one of knowing” (Smyth, 2001a, p. 629). It also rings true of Danziger’s suggestion that psychologists’ disciplinary discourse is actually the discourse of their investigative practice (1996). The structure of argumentation, the rhetoric the textbook author employs, is a model for a psychologist-to-be to learn how to argue her points in the future, be it as a researcher or as a professional. It teaches an interpretative and justified knowledge-making culture for producing psychological phenomena.

Here, the pedagogical expands into disciplinary formation. In the pedagogical practice of textbook writing, literally in how they were written, we can see some of the ramifications on what students do later as psychologists. Thus, the textbook is not only an implement to teach on a theoretical level, but a practical model-example.

Smyth ventures further in her analysis, trying to understand why the difference in the practice of writing textbooks between biology and psychology arose in the first place: "Psychology [. . .] presents evidence in its textbooks to override the engagement of the reader's everyday knowledge. There is no explicit, direct indication that folk psychology or common knowledge is to be replaced [. . .] yet the continual reference to evidence indicates that there is an argument with the reader going on, or the possibility of one. The reader is to be convinced that these accounts are reasonable, not told that they are so" (Smyth, 2001a, p. 632). The readers of a psychology textbook have direct access to their own psychological reality. They are directly at the source, and to be convinced that there is more to know about it than direct experience, something different than an autonomous fact statement is needed. An epistemologically superior method is required, and the hedging and the qualifications in textbooks show off precisely that. Echoing the Quine counterfactual read from Morawski's view on textbooks, the method in psychology is competing with common sense; it is not an extension of it.

This turns Latour's understanding of modalities, of hedging and qualifying, upside down. Modalities in psychology are employed to provide solidity for the claims, and in doing so, those very modalities are legitimized and unquestioned. The psychological claims in textbooks are a flood of evidence—they blunt opposition by empirical corroboration, a mob of studies and articles pregnant with experiments and sound research designs, signed with a name and a surname, and overtly supported by institutional and disciplinary affiliations. They are the heritage and the system empirical psychology leaves—the practice, not the knowing. And historians can explore the historicity of that practice through textbooks.

2.2.5 The virtues of the alternative view

The alternative view clearly provides a different way of historicizing textbooks in psychology. We can link it to how Kurt Danziger (1990, 1997) approached investigative practice — he extensively used journal articles to situate his history of psychology's investigative practice in order to learn more about the discipline. Another route to historicizing psychology can be traveled by looking at textbooks, providing a view of psychology from a counterpart to investigative practices: The practices of teaching psychology. As in the example above, maybe Danziger's investigative practices and the pedagogical practices turn out to be one and the same, the building blocks of psychological discourse in (at least) the second part of the twentieth century. There is a timeline to these pedagogical practices, their development and change, which can be tracked through textbooks.

Smyth's investigation of fact-making and Morawski's analysis of discourse in constructing subjectivities provide a framework for a diachronic perspective on textbooks. When did the practice of hedging and qualifying statements Smyth describes appear in textbooks? What was the referencing practice in the period Morawski investigates, at the turn of the century? In turn, what is the authorial voice in the later textbooks Smyth focuses on? These are all questions the historiography of psychological textbooks should engage with, but that have no place in the received view of most psychologists when they write about textbooks. With this short excursion into the work of Morawski and Smyth, we can easily recognize the constrained perspective the received view provides. It almost feels like it stops precisely where the fruitful analyses of the psychologist-historians begin.

2.3 Attempt at an integration as a conclusion

This chapter examined the received view of psychological textbooks in the journal *Teaching of Psychology*. I argued that this standard view is quite different than the trends in wider history of science when talking about science textbooks. The difference is not a consequence of the research object—psychological textbook versus other science textbooks—but rather a consequence of psychologists describing textbooks for an audience of psychologists, not historians. The psychologists doing this work were the authors publishing about introductory (and other) textbooks in *Teaching of Psychology*. What made the approach of these teachers of psychology into a history was Weiten and Wight's chapter *in Teaching of Psychology in America: A History*. The chapter recast the products of the teachers of psychology writing about textbooks into a history, transferring it from the context of pedagogical investigations of textbooks to a full-blown history of these textbooks. What was aimed to tell us more about textbook selection procedures by instructors, content representations and misrepresentations, and other points of interests of the teachers; suddenly became a story describing the change and role of textbooks through time. A history of this sort is a history told for the audience of those teachers in the first place, for whose research it serves as an overview.

The received view falls flat as a history in the professional sense when compared to the work of Morawski and Smyth precisely because these authors write for historians (or psychologist-historians). The rich descriptions of contexts be it social or intellectual or both, the massaging of the minutiae of particular discourses, the tracing of the construction of objects/subjects or viewpoints, the disciplinary negotiations, and descriptions of boundary-work are all blatantly missing from the received view.

More specifically, what is the difference between these two conceptions of textbooks, the alternative and the received? One is a site of fact-making versus an aggregator of facts; a place for constructing subjectivities versus just describing them; a context of a discipline versus a portrait of a discipline; a product engaged performatively with its context versus a text consisting of numeric-variabilized characteristics; a boundary object versus a manual of the discipline. It is relatively easy to argue that they are radically different and those differences are insurmountable—go about it as Leahy (2002) did in his scathing criticism of (a few particular) psychologists writing history. In an oversimplification: Psychologists write bad histories of their discipline, quantifying left and right, and we need historians to explain what is what in a nuanced and sensitive way.

I do not think the above is the case. Among other things, this chapter serves as an argument precisely against such scholastic distinctions—the citational analysis used to explore the work published in *Teaching of Psychology* is a case in point. I endorse the view that both the received view and the alternative need to be put side by side, or better said, step after step, in defining textbooks as historiographical objects.

In doing that, textbooks become historically constituted sites of fact making and constructing subjectivities, acting as boundary-objects between laymen and experts. Moreover, these historiographical objects are represented and described by actors functioning within the context, or a family-related one, those textbooks describe in the first place. The received view shows us that: How the discipline of psychology in the 1970s a historian would try to understand, for example, suddenly becomes the very part of textbook historiography through the work of Quereshi, mediated by Weiten and Wight. The psychologists' practices got extended into historical writing about psychology textbooks, collapsing the border between historiography and history. If we recognize the received view as a history of psychology, then our object is not just the textbook, but the textbook plus its surrounding historiography.

By conceptualizing it like that, we make evident the interplay between psychologists writing textbooks about the discipline that at the same time frames their way of thinking. One has to extend the circle to include psychologists writing about psychologists writing textbooks. A perspective like this integrates the received view of Weiten and Wight with the different readings aimed at other audiences than just psychologists. We can recognize their quantification of textbooks as one of the legitimate practices of the historian, but we cannot see it as the same thing historians do when they quantify—the quantification in Weiten and Wight's chapter is the consequence of a research practice being transplanted from the literature on teaching of psychology. There is no evaluative judgement here—I think there is no point in following Leahy in saying A is bad history, while B is good history (instead of good one could also say correct/truthful/objective and still miss the point). Both are histories aimed at different audiences, written with different goals in mind, and we should use both to come to a historiography of textbooks.

Integrating historiographies is not such a radical suggestion, especially if we consider it is not without precedent. On the contrary, in the larger context of history of psychology, the methodological bifurcation in textbook historiography between psychologists and psychologist-historians is the odd one out. The psychologists' toolkit — arithmetization in Danziger's words, the language of variables and statistics — already has a place as one of the ways of creating historical arguments. Danziger and Dzinas's (1997) investigation of psychology's metalanguage of variables is an example of such a content analysis (applying it to journals, not textbooks) used to make a historical claim. We can see similar methods used in Andrew Winston's (2004) investigation of the same metalanguage in textbooks. Extending it even further, large scale analyses of texts are a veritable research field in history of psychology, with pioneers such as Christopher Green (e.g., Green & Feinerer, 2015; Green, Feinerer, & Burman, 2015a). Obviously, the point I am trying to make is not an all-out rejection of quantitative analysis in history of psychology, or in the history of introductory textbooks in psychology.

It feels like the received view made the first step of generating content, but then never interpreted it historically. The textbooks just remained the very minimalistic description in the terms of the already established metalanguage of psychology. The force of the numbers was used to buttress a simple chronology, and not a contextualized and reflexive account of the role and function of textbooks in psychology.

One possible venue of integration is recasting my account of textbook historiographies into Latour's (2005) actor-network theory (ANT)—looking at textbooks as pieces in an expanding network of associations. Textbooks get transferred and transplanted from one context into another, reformed and recast with various meanings. Describing it in these terms, we could look at the integrated historiography being the result of three translations—the first when the textbook was taken from the classroom and the market by the authors publishing in *Teaching of Psychology* from the 1970s onward, who tried to optimize its selection and content with statistical methods. Then this expanding body of research consisting of hundreds of articles in our citation analysis were translated and synthesized into a history of those textbooks— the textbooks have turned from an optimizable and faulty receptacle of facts into a series of portraits of a discipline, effectively glimpses into the essential discipline of psychology. The last transfer was conducted in this article, where these portraits of a discipline crafted by Weiten and Weight were downplayed by describing their construction of subjectivities (through their content) and objectivity (through their fact-making techniques) endorsed by psychologist-historians.

Every translation has changed the textbooks—the first refashioning them into objects susceptible to statistics, the second using these quantitatively represented objects as windows into history, and the third designating these windows as a stepping stone for historical interpretation. Keeping with the Latourian allegory, the thousands of psychology textbooks sketchily marched through assemblages of psychologists and their statistical methods, psychologist-historians, historians, journals (and their audiences), editors, and edited volumes. The textbook that existed before authors in

Teaching of Psychology started writing about them, and the ones we have at the end of this chapter are almost nothing alike.

I would keep this ANT interpretation as an allegory that helps us understand the argument in this chapter, but not a definite conclusion. After all, particular textbooks were hardly dealt with in this chapter, and the intricacies of ANT's descriptions of actors and their movements through the network would needlessly overburden the whole argument. It also chronologically misrepresents the actual research on textbooks—what is called the received view still thrives in *Teaching of Psychology* as part of research on the classrooms where psychology is taught; the articles and the textbooks constructed in them are quite unscathed by what was called Weiten and Wight's translation in our allegory.

The thesis I sketched in the chapter's introduction—that the received view of textbooks in psychology is different from the view of textbooks in wider history of science still stands firm. The textbook as an object produced by the publications in *Teaching of Psychology* is something radically different than the object Morawski, Smyth, and Olesko describe in their work. However, my account of the received view was not produced to criticize this difference—agreeing with, for example, Morawski's view of the textbook at the expense of the one produced by Weiten. It was aimed to produce a textbook as a historiographical object that bridges this chasm between the standard view and its counterpart.

Casting it Hacking's terminology (2002), the historiography of psychological textbooks should follow their historical ontologies by traveling through various contexts—the ones they were written in (author's), the ones they were written for (classrooms), and the ones they were debated in (among psychologists and historians). The textbook, as a nexus of all kinds of forces—of the academic community (both of psychologists and historians), of the pedagogical necessity, marketing of the discipline's practical utility, authors idiosyncrasies, the interest of the publishers—is too fruitful to be left to just content analysis, or just discourse analysis.

Trying to find one last lesson from this search through the muddled and disconnected scholarship of textbooks, I would argue that this makes psychological textbooks enter center stage of historiography of science through the backdoor. From a marginal historical space confined to science pedagogy and book history, it offers itself instead as one of the possible platforms for debates of disciplinary formation, fragmentation, and identity. Textbooks become reifications of investigative practices coupled with pedagogical ones. They mesh the contexts of theoretical knowledge and application, the rhetoric of presenting it, the business of teaching it, and provide objective and objectified perspectives on a discipline, perspectives that can be carefully recast and inspected. They are full of "facts" that at the same time bring and hide uncertainties they are textual witnesses of a time gone by. Textbooks (again?) become a place to go to reflect about the discipline. A good Latourian might even say textbooks gain agency in driving that reflection.

Chapter 3. The stable core of an unfinished science

Introductions of undergraduate psychology textbooks are not the first source that comes to mind when one goes looking for good historical accounts of 20th century psychology. Historians after the cultural turn – the turn roughly identified in the discipline's historiography with "new history of psychology" (Furumoto, 1989)[53] – shy away from published sources that are obviously biased from a historical point of view. Following Furumoto and the new historians, textbooks seem like a stark example of a source that was not, in its time, produced to convey an accurate picture of the discipline. They were intended to provide a snapshot of what the authors thought the discipline ought to be, or at the very least what the students should think it was. Textbooks froze in time an image of the discipline that had to be comprehensive, simplifying, educating, persuasive, abstracted, etc.; all at the same time! None of these criteria necessarily ensured an accurate historical picture. Further muddling the issue, the textbook authors also molded their account into a personal view on the complicated disciplinary nature of psychology. The tension between all those incompatible objectives, which made textbooks both less and more than an accurate depiction of the discipline, was the consequence of the conflicting requirements set before the author by the different audiences s/he was trying to address.[54] The book had to speak to two very different publics: Naïve students and cautious fellow experts. And then, on top of it all, the final arbiter deciding whether the text was good enough to see the published light of day was the publisher. It follows then that even when particular textbooks were seen as intellectually worthy in their own time, they often didn't age well. They were a product torn to serve conflicting purposes, often succeeding at none.

Many scholars think that textbooks are of dubious intellectual worth, not only as objects of academic historical study but even as sources of information for students. Psychologists themselves have argued so on many occasions (Ferguson, Brown, & Torres, 2016; Griggs, 2014a, 2014b, 2015a, 2015b; Gliner, Leech, & Morgan, 2002), saying that textbooks are rife with faults and misconceptions. Most of the articles I just cited discuss distortions in the depiction of landmark studies: Little Albert, bystander apathy as linked to the murder of Kitty Genovese, and the Stanford Prison Experiment. If we are to believe the cited psychologists, textbook representations of psychological knowledge are untrustworthy. Among philosophers and historians, Kuhn's view

---

[53] Even though the program might not be new, the goals of "new history" are still a hot topic of discussion by historians of psychology; see Lovett (2006; 2017), Brock (2017a; 2017b), Watrin (2017), and Araujo (2017).

[54] For the molding influence of publics on psychological knowledge and its production, see Pettit and Young (2017).

(1962/2012) of "textbook science" distorting history is one of the most well-known.[55] Both psychologists and historians/philosophers seem to agree that the prospect for any use of textbooks as historical sources, after their prime when they had been used in classrooms, is even glummer.

I will take the contrarian position: I will use a set of textbooks as historical sources precisely because of their faulty, biased, and motivated character. My view is that they are a source of structured distortion that is historically relevant. Tracking what gets distorted and how lets us learn what were the contentious issues and insecurities of the time. In textbooks, we find what made the psychologists writing them tick – the good and the bad. The bad being the contemporary anxieties in the field; while the good were the authors' hopeful plans for the discipline's future. The textbook's story was never simple – it was an assemblage of wishful thinking of the author, comprehensive overviews of different research programs; and the inherited, deeply entrenched metaphysical positions about science and human psychology. By looking at fifty years of editions of a single textbook, I follow the continuities and changes in this story of psychology through time.

As my analysis will show, we can learn something interesting from carefully reading the editions of the same textbook spanning five decades: Psychologists have constructed a particular niche for their science which safeguarded their discipline both in the university environment and in the wider society that demanded applicable knowledge. According to the textbook authors, psychology was barely out of its infancy during the 20th century. By that century's end the discipline had a stable methodological core which held the promise of uncovering systematic and true psychological knowledge at some point in the future. Psychology was scientific in the sense of the received view that science was the kind of activity that accumulates knowledge and brings humans incrementally closer to the truth about the world. I will show in the article that psychology's niche – a rhetorical space, a territory carved from the subject matter of the empirical sciences and the university departmental structures - was maintained and informed in a consistent way since the 1950s, by defining the discipline as an unfinished scientific program within the larger project of science. The problem is, as I will argue in the chapter's conclusion, that this epistemological program may have put too much faith in methods alone to transform[56] psychology into something more than an unfinished science. I will argue

[55] For a historiographical overview on psychological textbooks, see Chapter 2. For a periodization of American psychology textbooks, see Weiten and Weight (1992). For relevant case studies that go beyond Kuhnian readings, see Winston (1990), Winston and Blais (1996), Morawski (1992; 1996), and Vicedo (2012).

[56] "Transform" and the vocabulary I use imply that I uncritically accept the actor's category of scientific progress. I do not, but that issue is too complicated to discuss at this point. For now, take my usage of the vocabulary of psychology's progress as adopting actor categories, not as my hope or conclusion. The otherwise historiographical problem is complicated by the fact that there is a close continuity between the

that psychology remained an unfinished science *because* of its methods, and not in spite of them. My analysis is historical and rhetorical, but my conclusion will be polemical vis-à-vis the contemporary debates among psychologists that are commonly subsumed under the label "replication crisis" (Green, 2018). I shall argue for my position in the rest of the chapter in three ways: By the choice of the textbook I analyzed (the what), the way I did it (the how), and the conclusions I have drawn from my investigation (the what for).

3.1 The what

I shall first turn to my choice of a particular textbook. I investigate a single influential book that went through thirteen editions in the second part of the 20th century: Hilgard's *Introduction to Psychology*. Ernest "Jack" Hilgard was an important member of psychology's pre- and post-World War II elite at American universities. As Kihlstrom put it in an *In memoriam*: "Jack Hilgard lived the history of psychology in the 20th century. He met Pavlov, argued with Skinner, and nurtured many of the first generation of cognitive psychologists at Stanford. All of his work was informed by a consciousness of the past" (Kihlstrom, 2002, p. 97). His undergraduate textbook was part and parcel of his status as one of America's Great Men in psychology: "Hilgard had a real talent for expository writing [...] *Introduction to Psychology* [...] was by far the most popular introductory textbook of its time and set the standard by which all other introductory texts are judged. In addition to presenting the fundamental concepts, principles, and methods of scientific psychology, the introductory text indulged Hilgard's proclivity and talent for 'psychologizing'" (Kihlstrom, 2002, p. 96-97). The book was not only "one of the most widely used books in the history of college publishing" (Atkinson, Atkinson, Smith, Bem, & Hilgard, 1990, p. V), but by the year 2000, had been translated into French, German, Hebrew, Italian, Portuguese and Chinese (Atkinson, Atkinson, Smith, Bem, & Hoeksema, 2000, p. V).[57]

The textbook's later editions were written by Hilgard and a number of co-authors: The 4th edition (1967) with Richard Atkinson; the 5th (1971), 6th (1975), 7th (1979), and 8th (1983) edition with Richard Atkinson and Rita Aktinson; the 9th (1987) with the Atkinsons and Edward Smith; the 10th (1990) with the Atkinsons, Edward Smith, and

---

history of the discipline in the 20th century and contemporary debates – historical actor's categories are coextensive with contemporary actor's categories.

[57] The observation that it was one of the most influential books in the history of college publishing was taken from the preface of the textbook's tenth edition, so we should take it with a pinch of salt. However, the book's importance and widespread use is beyond doubt. For example, it was perceived as the dominant textbook in some parts of Europe (Newstead & Makinen, 1997, p. 8). Hilgard's textbook can be looked at not only as a cultural artifact for disciplinary indoctrination of students in America, but also as a tool used for the spread of American psychology internationally, which, according to some historical reconstructions, had explicitly political overtones during the Cold War (for an example in social psychology that is unrelated to textbooks, see Schruijer, 2012 and Schruijer, 2008).

Daryl Bem. The last three editions of the 20th century were co-authored by the Atkinsons, Edward Smith, Daryl Bem; with the addition of Susan Nolen-Hoeksma for the 12th and the 13th edition and of Hilgard's surname to the title of the book, making it into *Hilgard's Introduction to Psychology*. The social history of how historical actors wrote and negotiated the book, and how it was received by psychologists and specifically influenced individuals, although extremely interesting, is not the topic of this chapter.[58] What I aim to do here is a chronological rhetorical analysis of the disciplinary overviews of psychology presented in the textbooks' introductory chapters. The fact that the book was produced by the members of the elite centers of American psychology and that it was continuously published through the period of fifty years makes it a perfect candidate for such an analysis. On top of all the unique features a textbook story provides, *this* textbook story was also extremely successful, long-lived, and widely circulated.

3. 2 The how

The way I analyze the editions is by closely reading the chapters that provide a broad overview of psychology. In most editions those are the introductory chapters.[59] There the authors give a broad-strokes picture of what is psychology. Giving a short overview of something as big and disjoined as psychology requires a lot of rhetorical work from the author. Rhetorical "in the neutral sense, as an inherent dimension of any form of communication" (Sommerey, 2015, p. 14; for more on rhetoric of science, see Gross, 1990; Nelson, Megill, & McCloskey, 1987). I aim to illustrate the chronological changes and continuities in the textbook authors' rhetorical strategies to persuasively give their view on what psychology as a discipline was. These rhetorical strategies produced arguments and evaluations of the field which were general and ideal-typical. They were the views of the author, an insider motivated to "give psychology away" (Miller, 1969; for a discussion on "giving psychology away" in the case of textbooks, see Morawski, 1992). Even though these are the authors' generalist perspectives, there is an expectation that the book changes through time – especially when we think of all the requirements put in front of a book of this nature by the changing academic context, the developments in psychology, and wider society. On top of all that, different editions were written by different authors.

---

[58] For a historiographical discussion of the influence of textbooks on psychologists' views of their own discipline, see Chapter 2 and Winston's analysis of Woodworth's "Columbia Bible" (1990). The view that textbooks are vehicles for socializing identities of future professionals is still an under-researched subject in the history of psychology.

[59] In the first edition, the broad topics are covered in the first and the last chapter. Through the editions, the various important points informing broad overviews of psychology move about the book, from the last chapter to the first, or from the first chapter to the appendixes at the very end of the book. When such movements happen, I make a note of it in my analysis.

Textbooks don't just change through iterations of editions. Certain parts of the book stay the same, since supposedly it is the same book being republished. The authors themselves were aware of this, when in the preface of the fifth edition in 1971 they wrote:

> There is an old story concerning a peasant housewife whose blanket kept unraveling at one end, and who kept knitting on an equivalent amount at the other end. After all the material had changed, was it the same blanket? A similar question may be asked of a textbook going into a fifth edition, with each edition thoroughly revised. Is it still the same textbook? The reply in each case is a conditional "yes," for both blanket and book serve the same purposes today that they served in the past, and there is continuity in the midst of change. (HIP5: V)[60]

Following the authors' metaphor, I look closer at the particular strands woven into the blanket. By analyzing the text itself, I try to uncover the relational character[61] of the textbook's introduction that makes it unique as a historical source: It's a synthetic evaluation of the whole field produced by an elite insider that's simple enough to be understandable to students, but more or less agreed upon by the community of psychologists the author is trying to represent. I say "more or less" to avoid over-interpretation. Psychologists in both the first and the second half of the 20th century were a quarrelsome bunch. Hilgard was very much aware of that, trying to accommodate most of them with his textbooks. In the preface of the second edition (HIP2: V), he wrote:

> Because I sought to make the findings of psychology both clear and plausible, without distracting the student too much with the quarrels of psychologists, in some parts of the first edition I invited the charge of dogmatism, of placing in an awkward position those instructors with strong convictions that did not agree with mine. In this edition I have taken care to call attention more positively to unresolved issues. Discussion of these unsettled problems encourages the students to do their own thinking and sustain the picture of psychology as an unfinished science.

---

[60] A note on in-text citations of the textbooks under analysis: From hereafter, in order to streamline and make the chapter more readable, when citing from Hilgard's textbooks I will use the following format – HIP5: V, meaning the cited passage is from Hilgard's *Introduction to Psychology* edition 5, page V.

[61] By relational I mean it in two ways, both contemporary for the textbook author(s). In the sociological sense, it tries to connect the psychologist's expert view to the layman's commonsensical view; and in the content of the science presented, it tries to connect a broad and simplified perspective of the discipline to the many differing research fronts the researchers in psychology are engaged with at the time. For a sociological Latourian analysis of psychology textbooks, see Smyth (2001, 2004a, 2004b). For a large-scale analysis of research fronts in psychology in the period from 1950 to 1999, see Chapter 4.

Since the same text is rewritten and reused from edition to edition, I adopt what I call a modular view. Fragments of the text from the previous editions are sometimes kept verbatim and just rearranged in the later editions. In other places they are edited, trimmed, or expanded. I call these fragments modules because when followed through a number of editions they are reconfigured to form compact parts of the introduction. These compact parts serve identifiable purposes. Considering this, I will discuss the following groups of modules:

a) The opening paragraphs of the introduction. I call this group of modules *the pitch*.
b) The bulk of the text in the introductory chapter which I label as *the discipline overview*.
c) The part of the introduction that gives students insight into the methods of psychology, which I call *methods*.

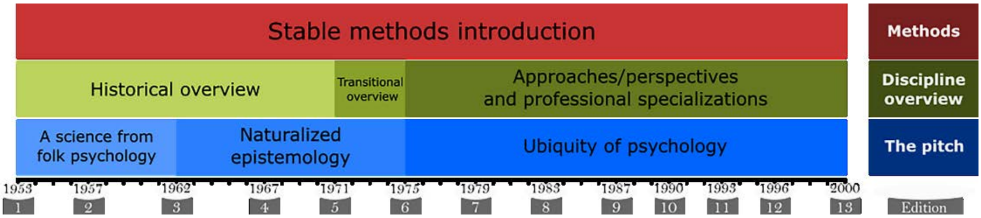


Figure 3.1. Schematic of the thematic analysis of textbook introductions

A schematic view of my rhetorical analysis through the thirteen editions of the book is presented in *Figure 3.1*, showing how the approach to the pitch of the introductory chapter and the overview of the discipline changed, while the introduction of methods stayed the same. The horizontal axis shows the publication year of the editions; the number below it is the edition number. The bottom blue band represents the thematic changes in the pitch modules. In the middle, the green band represents the changes in the modules related to the bulk of the introduction. The top red band represents the stable modules which introduced methods of research. I recommend the reader to use this figure as an organizing blueprint for following my analysis in the chapter, to which I turn next.

3.2.1 The pitch

The pitch consists of the opening paragraphs of the book's first chapter. In it the textbook authors attempt to draw the reader in. If we look at the textbook as a consistent treatise introducing an area of scientific knowledge, the pitch acts as a hook for the whole exposition. It sets the tone of what is to follow. For Hilgard's textbook, it goes through three large shifts since the first edition in 1953 to the thirteenth in 2000.

3.2.1.1 Building a science from folk psychology

The first edition introduced psychology as a subject and a science that was quite relevant for everyday life. The author reminded the student that "[o]ur folklore is full of psychological statements: 'As the twig is bent the tree's inclined,' 'Practice makes perfect'…", and that thinking as a psychologist is a part of everyday life: "As we form theories of human conduct, each of us becomes in some real sense his own psychologist" (HIP1: 4). Immediately afterwards, under the heading "What to expect from psychology?" the author wondered if the student will find answers to the questions that might interest him during a course in psychology: "Will he be able to form better judgments of other people? Will he be able to plan more wisely for himself? Will he make friends more easily?" The psychologist's answer was, as the authors put it, "both 'Yes' and 'No'" (HIP1: 4). As Hilgard went on to argue, although psychology was relevant for every-day life, it was not the type of knowledge one found in self-help manuals: "A course in psychology is not a course in self-help, and a psychology textbook is not a manual on the art of handling people." (HIP1: 4). Such a pitch module, which built a science from folk psychology, appeared in the first and second edition.

An everyday-life approach is a common discursive tactic for psychologists writing textbooks (discussed in more detail in Chapter 2), the tactic of which Hilgard's first edition was a forerunner (Weiten and Wight, 1992, p. 467). Textbook authors wanted to make their discipline relevant to students – that's what made it interesting, and as the textbook authors have mentioned in many of their prefaces, they wanted to engage students on the students' own terms. But at the same time, while doing so, the authors ran the risk of bastardizing the science, diluting it into something self-evident and banal; something uncomfortably similar to what trained psychologists perceived as self-help. Striking the right balance was a delicate matter for the writer, oscillating between commonsensical obviousness and esoteric tedium of psychological laboratory facts. In essence, the authors had a nuanced message for their students: You ought to be interested and intrinsically motivated for learning about psychological facts, but *only* facts that were constructed scientifically.

3.2.1.2 Naturalized epistemology

In the third edition published in 1962, the pitch of the textbook moved from commonsensical folk psychology to the rhetorical strategy I labelled as 'naturalized epistemology.' Naturalized epistemology is a controversial position in 20th century philosophy of science (for an overview, see Rysiew, 2017). I do not wish to add to the philosophical discussions about naturalism in epistemology, but I wish to use naturalized epistemology as a name for the specific way Hilgard described the relationship between psychological science, psychological knowledge, and the natural world. By calling Hilgard's view naturalized epistemology, I want to illustrate how a nuanced metaphysical position was a crucial component of introducing psychology to students.[62]

The textbook's opening sentence of the first chapter in the third edition posited the need for understanding as one of the fundamental features of human mental life: "Because he can reflect upon the past, take account of present experiences, and make plans for the future, man has always sought to understand himself and the world about him" (HIP3: 2). Out of this need to understand, humans started to organize their impressions of the world into comprehensive systems. At first the comprehensive systems were religion and magical thinking, which were, according to Hilgard, obviously unscientific. But, "[t]he roots of science began when he [man] started to find some sort of order in natural occurrences that made them comprehensible;" Hilgard continued, "when, for example, he found that he could control his food supply by domesticating animals or planting crops" (HIP3: 2). Thinking scientifically was a natural state for humans, because "[s]cience leads to understanding of natural events; it leads to predictions about their course and therefore to some control over what takes place" (HIP3: 2). Psychology was an extension of this scientific approach, because it sought "to comprehend, to predict, and to control, taking as its special subject matter the behavior of man and the lower animals" (HIP3: 2). The natural need of humans to understand the world around them used to find its expression in religious or magical thinking, which was then slowly replaced by science. The scientific object of interest in psychology – the behavior and the mental life of individuals – followed a similar development, and because of that, psychology was a logical expression of science looking inwards instead of outwards. Psychology was a garden-variety science with a specific subject matter.

There is circularity in Hilgard's naturalized epistemology. Humans by nature are scientific. Part of that scientific outlook is looking inwards, seeking to explain one's own psychology. In other words, by working within the framework of psychology as a science we explain human nature, the very nature that drives us to adopt a scientific outlook![63] Hilgard seemed to find such circularity a useful tool for introducing his way of thinking to students. Psychology gained double legitimacy – as a science that spelled out what it meant to think and act scientifically, but also as a crucial part of the larger project of science which investigated humans as a part of the natural world. It also, at least indirectly, meant that psychology was a natural science.

Another thing to note about viewing psychology as informing and being informed by naturalized epistemology is how important it was for the textbook author to

[62] A more in-depth discussion of naturalized epistemology will be provided in Chapter 5.

[63] Many historians of psychology engage with the reflexive relationship between the subject matter of psychology (lowercase 'p' psychology) and the discipline which produces knowledge about that subject matter (uppercase 'P' Psychology; Richards, 1987, p. 204; 2002). For an anti-reductionist historical investigation of reflexivity in psychology, see Smith (2007).

communicate that psychology was a younger sibling of the natural sciences. In the second edition, the author puts it like this: “Psychology, as a behavioral science, touches and overlaps the other behavioral sciences. But psychology leans also toward physiology and toward the physical sciences. […] Psychology in relation *to physical sciences* is chiefly a borrower” (HIP2: 5, emphasis in the original). This kind of explicit boundary-work was present in all the introductions from the 1950s to the 2000s, later on demarcating psychology as one of the behavioral sciences among the social sciences, e.g. the two paragraphs on “behavioral and social sciences” in the eight edition (HIP8: 18).[64]

After the pitch with naturalized epistemology, the self-help module from the previous editions followed. The author’s innovation in it was touching on the incompleteness of psychology as a science:

> Psychological science, furthermore, is in an early stage of development. Many of its facts are not yet firmly established, and theoretical interpretations are often controversial. Sometimes we shall be studying more about how psychologists seek answers than about the results they have found. (HIP3: 2)

Taking a step from naturalized epistemology to the incomplete science conclusion was a straightforward move – psychology’s role as a science was reciprocally contained: It was both a subject matter and a collection of methods for producing knowledge about that subject matter.[65] For mainstream American psychologists at the time, psychology was a natural science, but a young one that was still experiencing growing pains.

The view that psychology was an incomplete science appears in various formulations throughout the editions of Hilgard’s textbook. One of the most explicit descriptions, which exactly echoed Mary Smyth’s Latourian analyses of psychology textbooks (2001, 2004a, 2004b) can be found in the textbook’s preface in 1971:

> We have retained the explicit documentation of statements made in the text through citation of sources, because we believe that psychology has not yet reached that stage in which declarative statements can be made dogmatically and anonymously. Most assertions have to be qualified by the context in which they have arisen, and this context is often best indicated by citing a source. Statements such as “Emotions disrupt learning,” or “Women change their attitudes more readily than men,” may be true in some contexts, but if one questions the statement, he should be able to find where it originated, who asserted it, and what the circumstances were. (HIP5: vii)

[64] For the politically controversial origin of the label “behavioral sciences”, see Pooley (2016).

[65] For “reciprocal containment” in naturalized epistemology, see Quine (1969: 83).

Central for Smyth's argument is the discussion of referencing practices in textbooks. Hilgard's textbooks conform to Smyth's conclusions: In-text references are present from the first edition in the form of footnotes. By the textbook's third edition in 1962, the *APA's Publication Manual* (for a history, see Sigal & Pettit, 2012; for a rhetorical analysis, see Bazerman, 1988, p. 257-277) exerted its full influence over the referencing style, with the standard form of surname/year in parenthesis used throughout the book. Hilgard included a footnote the first time such a reference was used, aimed at clarifying to the student what was the purpose of this reference: "Throughout this book you will find references to studies which document or expand on the statements made here. Detailed bibliographical information on these studies appears in the list at the end of the book" (HIP3: 3).

Smyth concludes the same in her comparative analysis of textbooks in psychology, biology, and statistics: Psychology was an uncertain science. As Smyth puts it: "Psychology textbooks do not present certainty, they present evidence" (2001, p. 411). Hilgard's comment in the preface corroborates Smyth's interpretation. The uncertain style was an explicit choice of the professional community – a choice that textbook authors wished to inculcate in students. Psychology was an organized way for reducing uncertainty by producing and presenting evidence, it was not a stable and abstracted system of evidence and generalizations structured into comprehensive theory. What was necessarily structured was the investigative process, not the theoretical claims produced by those investigations.

The naturalized epistemology module of the pitch appeared in the third, fourth, and fifth editions with slight variations. For example, in 1971 the hierarchy of sciences from hard to soft was included, considering the difference in the object of research:

> The physical and biological sciences were the first to be developed because the basic concept of orderliness is readily observed in the movements of the stars, the turning of the seasons, and the cyclical changes in trees and plants. [...] This capacity to keep some distance from the facts of observation – not to be too much influenced by personal preferences or prejudices – is the essence of objectivity, the dispassionate search for understanding that science embodies. (HIP5: 2-3)

What Hilgard perceives as "the essence of objectivity" is exactly the methodological imposition that produces the norms of his uncertain science. More on that in the chapter's conclusion. By the sixth edition in 1975, the pitch module of the book changed to the form it kept for the following 25 years until the thirteenth edition in 2000.

3.2.1.3 Ubiquity of psychology

In the middle of the seventies, the academician's naturalized epistemology gave way to the vision of a thoroughly psychologized society. The pitch modules showcasing the sheer ubiquity of psychology in society were the opening of the textbook from 1975 to 2000. I will analyze them on the example of the sixth edition. The sixth edition's opening sentence was: "Psychology touches almost every facet of our lives" (HIP6: 4). The need to legitimize psychology from the get-go was absent because psychology was everywhere. What *was* needed, though, was structure and an instruction manual to interpret all the psychological facts students were inundated with in everyday life. For this, they could turn to psychology. The textbooks authors continued:

> As society has become progressively more complex, psychology has assumed an increasingly important role in solving human problems. Psychologists are concerned with an astonishing variety of problems. Some are specific and practical. What is the best treatment for drug addiction or obesity? How should a survey be designed and administered to measure public opinion accurately? How can people be persuaded to give up smoking? What is the most effective method for teaching children to read? How should the dials on the instrument panel of a jet aircraft be arranged to minimize pilot error? Can a blind person be given artificial sight by electrical stimulation of small wires implanted in the brain? (HIP6: 4)

The barrage of disconnected questions cued the student to all the varied manifestations of psychology throughout society. Psychology was everywhere and touched on all kinds of very relevant or even explosive questions. For example, the authors discussed "the effect of television violence on children" (HIP6: 4). Psychological research being a constant presence in society also meant many claims were produced as scientifically legitimate. In order to be able to separate the informational chaff from the wheat, the students needed to receive at least basic training in psychology. Even if they would not become psychologists themselves, an introductory course in psychology "should also help you evaluate the many claims made in the name of psychology" (HIP6: 4). The need to be able to evaluate claims competently was illustrated by a list of newspaper headlines with such claims, for example: "Homosexuality linked to hormone levels" or "Proof of mental telepathy found" (HIP6: 4). How would the students, as consumers of news, have decided on the verisimilitude of such statements? The authors offered remedy with the introduction of scientific methods.

With their turn to methods in the pitch, the authors kept the idea of psychology as an immature science. The argument went like this: In order to be able to evaluate psychological claims encountered in the day-to-day, the student would have to learn how psychological claims were made in the first place. "How can you judge the validity of such claims? In part, by knowing what psychological facts have been firmly established and by being familiar with the kind of evidence necessary to give credence to a new 'discovery'" (HIP6: 4). The shudder quotes around 'discovery' reinforced the notion that psychology wasn't making discoveries; it made collections of rigorously constructed facts. As a lesson in the program of reduction of uncertainty, the student was taught about the way the psychologist conducted her research: "It [the textbook] also examines the nature of research – how a psychologist formulates a hypothesis and designs a procedure to prove or disprove it” (HIP6: 4).

The authors went on to say that “the recent years have seen a virtual explosion of psychological research” although “psychology is relatively young compared to other scientific disciplines” (HIP6: 4). As a result of this expansion of research literature, “psychological theories and concepts have been continuously evolving and changing” and this makes it “difficult to give a precise definition of psychology” (HIP6: 4).[66] In case of an unfinished science like psychology, it was much more prudent to provide a collection of facts and the rules for their production and testing. Students, as educated and competent consumers of information in the psychologized society, needed to be armed with the minimal skillset for judging fact from fiction. Of course, under the assumption that they hoped to increase their chances of navigating society successfully.

Incidentally, presenting psychology as an unfished science created the air of an exciting academic frontier. In an incomplete science there was still much room to expand knowledge, for a young individual to become a pioneer participating in a new and rewarding extension of scientific thinking. For the discipline, legitimacy was gained through its exciting newness and not its history. The quagmire of unresolved metaphysical and epistemological issues could be left to psychology’s philosophical past. This is not to say that the textbook did not provide a certain interpretation of the discipline’s history. To that group of rhetorical modules – the one concerned with overviews, historical and otherwise - I will turn next.

3.2.2 Discipline overview

Older editions of Hilgard’s textbook included large historical overviews in the introduction. This was the case in all of the editions in the 1950s and 1960s, except for the first one. In the first edition, the historical overview of psychology was given in the last part of the book, under the chapter heading *Psychology as a Science and as a Profession*. In the following editions this historical overview moved from the back of the book to the front. Despite the move, the modules carrying the historical introduction in the first and the second edition stayed largely the same.

The historical overview followed the general sense that Thomas Kuhn (1962/2012, p. 136) called a “substitute” history that can be found in science textbooks. Kuhn called this kind of account a substitute because it replaced the mess of science’s historical record with whatever would have fitted the reigning paradigmatic perspective. His

[66] The difficulty in providing definitions of psychology resulted in a fascinating exercise in organized uncertainty. The editions from 1970s to 1980s had a table listing a dozen or so definitions of psychology through the centuries. Depending on the edition, the listed definitions were from Great Men such as William James and Kurt Koffka, or in later editions by Kenneth Clark and George Miller (HIP6: 12; HIP7: 12; HIP8: 14; HIP9: 13).

view on science textbooks, though, will only get us so far. For Kuhn, the bits of history included in textbooks were there to provide a story that delimited the relevant problems and methods within the currently ruling paradigm. At the same time, they provided the current paradigm with valid reinterpretations of the works of Great Men who were seen as the paradigm's progenitors. Hilgard and his co-authors achieved a similar goal with their historical overviews, but the consensus about psychological theories was a much more tenuous thing.

When I say tenuous, I mean that the exposition on history in Hilgard's textbooks was tempered with a view on psychological theory. History was conjoined with an explicit discussion of theorizing, which I would argue is quite different from Kuhn's contested notion of scientific paradigms and their ersatz histories aimed at smoothening out the rough edges and educating 'normal' scientists. Tangentially, I would argue that the source of much worry for psychologists trying to learn something about their own discipline from Kuhn was that it is still unclear if psychology ever had a paradigm (for elaborate explorations of psychology's unity by psychologists, see Staats, 1983; Sternberg 2005) – whatever a paradigm might be.

3.2.2.1 History introductions in the 1950s and 1960s

A recognizable collection of modules introducing history and theories of psychology appeared in the first four editions of the textbooks. In these modules, the textbook author was of the firm opinion that introducing a science did not only entail empirical facts and their modes of production, but also theories. For example, in the 1953 edition, the history and theories were introduced with the following paragraph:

> Science is not merely an accurate and quantitative description of events; it is also an attempt at economical and systematic handling of its data in order to establish from hypotheses and theories verifiable laws. In this chapter we wish to consider some of the problems confronting psychologists as they seek to make a science of psychology. (HIP1: 550)

The message here was that the science of psychology was still a thing-in-the-making, and psychologists were at the frontier of extending scientific thinking to human affairs. As was reasonable for a discipline supposedly being turned into a science in front of the students' eyes, all four editions in the fifties and the sixties included a paragraph on what was perceived as disciplinary prehistory. It covered the Greeks, St. Augustine, and Descartes, and "many prominent philosophers of the seventeenth and eighteenth centuries – Leibnitz, Hobbes, Locke, Hume" (HIP3: 14; HIP4: 13).

What followed the prehistory was a short exposition on faculty and association psychology as the two relevant approaches in the 19th century. Psychophysics, with Weber and Fecher, made its appearance in two editions (HIP1: 551; HIP2: 7), with Wundt occupying a prominent place with a subheading "Wundt's laboratory" in all four (HIP1: 552; HIP2: 7; HIP3: 14; HIP4: 14), proclaiming the commonly held view of Wundt as the founder of modern experimental psychology. The importance of Wundt for American psychology was explicitly addressed in all four editions, considering "so many pioneers in American psychology went there [to Leipzig] to study" (HIP1: 552; HIP2: 8; HIP3: 15; HIP4: 16), mentioning by name William James, G. Stanley Hall, and J. McKeen Cattell. This coverage of history strongly echoed Boring's influential view on the history of experimental psychology (1929; 1950).The historical overview of the first four editions wrapped up with other sources of influence, bringing them to the student in the form of short snippets elaborating the contributions of Great Men: Francis Galton, Charles Darwin, Anton Mesmer, and Sigmund Freud.

The glimpses of history were focused on listing the achievements of Great Men and the cumulative development of psychology as a science. A great example for this style of writing history was the author's view on faculty and association psychology (all four editions contained a variant of this passage):

> Both faculty and association psychology have their counterparts at the present time, but with notable differences between the old and the new. The search for primary abilities underlying scores on psychological tests, which we will meet later, is related to faculty psychology, but it differs in its careful quantitative approach. Much of learning theory, especially the theory of conditioned responses, is similar to earlier association theory, except that now we believe that stimuli and responses rather than ideas are associated. Very often, thinking men of earlier centuries anticipated later developments. (HIP3: 14)

The short Boringesque history of psychology is nothing unexpected – it is a known feature of 20th century American psychology and its origin stories. The modules that followed the historical account and covered "the role of theory in psychology" (HIP3: 16; HIP4: 17) are much more interesting as the authors' views on the state of their discipline.

3.2.2.2 Theory introductions in the 1950s and 1960s

Introductions of history going hand in hand with discussions of theory were a staple for teaching psychology in America, usually under the heading "history and systems" (Hilgard, Leary, & McGuire, 1991, p. 95). The period of schools in the first decades of the 20th century has been built into the disciplinary conscience of American psychologists thereafter. What followed the schools period, in the received account of the discipline's post-WWII history, was a period of more lax systems:

> The period of schools in psychology is passing, but these systemic schools have served their purposes in providing rallying points for enthusiastic workers, in correcting faulty emphases within opposing schools, and giving some measure of unity to the complex fragments of psychology, even though the unity achieved may in some instances have been ill-founded. (HIP3: 16)

In the first edition, the argument for introducing competing systems of psychology, namely introspectionism, behaviorism, gestalt psychology, psychoanalysis, and functionalism was quite elaborate:

> With different theories as starting points, different facts emerge. It is possible to have two sets of natural laws, both true. They will not really contradict each other, but they may be so different as to make comparison difficult. Hence natural laws are not there in nature merely waiting to be discovered. They are the result of a complex process by which scientists select facts to gather and invent theories to fit the facts so collected. All this sounds rather abstract, and makes sciences seem less plain and businesslike than they usually seem. It would not be important to mention these peculiarities of sciences and of their laws were it not for the fact that there are competing theories within psychology. (HIP1: 555)

Competing systems implied that more than one theory about the same subject within psychology was true, which sounded like a logical impossibility. The author conceded the point, but with the caveat: "[I]t is possible that all of the theories are true, but limited. That is, they give correct answers to many of the questions that they ask; but all the systems do not ask the same questions" (HIP1: 555). And again, the conclusion was a hope for psychology as a new scientific frontier: "Perhaps ultimately one way of stating psychological principles will be found more useful than any other, but that point has not yet been reached" (HIP1: 555). The state of the theories in the field was explained by the notion that psychology was an unfinished science.

The three systematic positions introduced in the third and fourth edition were behaviorism and S-R psychology, Gestalt psychology (combined with field theory in the third edition; and cognitive theory in the fourth edition) and psychoanalysis. Behaviorism was introduced as an orthodox school of scientific psychology earlier in the century, which broke away from introspectionism and made psychology more scientific through the work of psychologists like John B. Watson. For example, the author wrote: "In order to make psychology a science, Watson said, its data must be open to public inspection like the data of any other science. [..] Behavior is public; consciousness is private. Science should deal with public facts" (HIP3: 17). Stimulus-response psychology was the contemporary and more relaxed heir to dogmatist behaviorism: "There are still a few ardent behaviorists, but most contemporary psychologists are not extreme about it" (HIP3: 17). The difference between S-R psychology and earlier behaviorism were covered briefly, with a particularly interesting comment in the critical discussion box that appeared in both the third and the fourth editions, describing how S-R psychology could potentially transcend a theoretical position:

> If very broad definitions are used, so that stimulus refers to a whole class of antecedent conditions, and response to a whole class of outcomes in the way of movements and products of behavior, then S-R psychology becomes merely a psychology of independent and dependent variables. [...] Viewed in this way,

> S-R psychology is not a particular set of theory, but rather a language which can be used to make psychological information clear and communicable (e.g. Mandler and Kessen, 1959). As such, the S-R outlook is widely prevalent in American psychology today. (HIP3: 19).

Similar views were expressed in the second edition (HIP2: 21). In the conclusion, I will return to this view of S-R psychology as a prototypical strategy for connecting the view of an unfinished science with methodological rigor by the end of the 20th century. S-R psychology might have died out as a dominant theoretical orientation, but its powerful rhetorical strategy lived on.

Psychologists' skepticism toward big systematic theories was reinforced in one more way in the introductions of the early editions – by arguing that contemporary psychologists subscribed to particular scientific models, not largescale theories: "Many psychologists refuse to give their loyalty to any closed or final system while the data of their science are being constantly revised, and while many relationships have not yet been satisfactorily studied" (HIP3: 21). Such more particular systems attempted to propose smaller "theories of forgetting, or theories of attitude formation, or theories of hearing" (HIP3: 21). Building smaller systems had a precedent in the physical sciences – psychologists even explicitly took such models from certain physical sciences as inspiration, with the authors giving examples of "atomic chemistry" and "field physics" (HIP3: 21). More importantly, building miniature systems was just one step in the direction of psychology's progressive development as a science: "The miniature system or more limited model saves science from becoming an unwieldy mass of scattered facts without forcing it prematurely into a mold that might warp its development. Until many smaller systems are securely established, a comprehensive system of psychology may be some distance away" (HIP4: 22). In other words, largescale theorizing could wait until enough facts had been amassed. Such an agnosticism about theories easily segued into an all-out anti-theoretical stance in the name of methodological rigor. Theories were a mold which warped perspective. Psychologists ought to have been producing data as evidence instead.

3.2.2.3 Approaches to psychology from 1970s to 1990s

Starting with the fifth edition in 1971, there was a break with how psychology was overviewed in the introductions. The fifth edition offered a sort of in-transition version of the introduction, from the systems and history one in the 1960s to a one focused on what the authors called the approaches to psychology in the later decades.

The new introduction started with a strongly biological exposition: "Psychology, as we know it, is a post-Darwinian science" (HIP5: 3). The authors then ran the student through a brief sketch of psychology's history by focusing first on behavior by covering behaviorism, S-R psychology, comparative (animal) psychology (HIP5: 3-5); then on conscious processes, awareness and mental activity by discussing Tolman, cognitive psychology, and introspection (HIP5: 5-7); finishing with the unconscious processes by briefly discussing Freud (HIP5: 7) and the role of the brain (HIP5: 7 – 8). This strategy resulted in a more convergent narrative – by discussing psychology through an evolving subject matter and the perspectives that have developed to account for it, the authors provided a compact coverage of the history which led to what were two relevant perspectives – S-R and cognitive psychology on one hand and what the authors call the "biobehavioral sciences" (HIP5: 7) on the other.

What shifted from the last section of the textbook to the introduction (like the historical introduction shifted in the second edition) was the part shortly describing the various specializations of actual psychologists; giving the data on the number of psychologists specializing in each in the US. Here, for example, the students learned that 48% of psychologists in 1969 were in the clinical and counseling/guidance specialization (HIP5: 11). Each of the larger specializations received a paragraph describing them - experimental psychologists; clinical and counseling psychologists; developmental, personality and social psychologists; industrial and managerial psychologists; educational and school psychologists; and methodologists. The discipline's overview in the fifth edition, then, consisted of the coverage of a progressively evolving subject matter and the collection of specialized professionals and "what they actually do" (HIP5: 10).

The editions from sixth to twelfth completely dropped the already trimmed historical introduction of the fifth edition and instead the bulk of the intro text was taken by what the authors called "approaches to psychology."[67] Fragments of historical information weren't completely omitted: The "behavioral approach" still mentioned Watson and his campaign for behaviorist psychology and on the other hand the "psychoanalytic approach" still discussed Freud. But a dedicated historical overview was dropped. The "approaches to psychology" took over as an organizing analytical category for the introduction. Instead of a development through time, one would read a development of subject matter.

Approaches in the sixth edition were described as types of truthful descriptions of phenomena that put emphasis on different parts of the situation. The authors provided the following example for the student:

> Any action a person takes can be described or explained from several different points of view. Suppose, for example, you walk across the street. This act can be described in terms of the firing of the nerves that activate the muscles that move the legs that transport you across the street. It can also be described without reference to anything within the body; the green light is a stimulus to which you respond by crossing the street. Or your action might be

[67] A separate coverage of the history returns to the textbook in the later editions as an appendix (e.g. "Brief History of Psychology" in HIP9: 631-639; HIP10: 765-773). In content, it's very similar to the history covered in the older editions, but it's included as one of the appendixes. In the thirteenth edition in 2000, a brief history is again included in the introduction chapter.

> explained in terms of its ultimate purpose or goal: you plan to visit a friend and crossing the street is one of many acts involved in carrying out the plan. (HIP6: 5)

In the same way that there were multiple ways to describe someone crossing a street, the authors concluded, "there are also different approaches to the psychological study of man" (HIP6: 5). In the sixth edition, the authors introduced psychology by describing five approaches on seven pages of the introduction. The approaches described were neurobiological, behavioral, cognitive, psychoanalytic, and humanistic. The same approaches were covered in the editions seven through thirteen – the cosmetic difference was that they were renamed into perspectives, and the humanistic approach/perspective was first renamed into 'phenomenological or humanistic approach' (HIP7: 8), and then into phenomenological perspective in the following editions. Later editions also complemented the main approaches with what was introduced as the "frontiers of psychological research" (HIP13: 16), which in the thirteenth edition included short paragraphs in a side-box covering cognitive neuroscience, evolutionary psychology, cognitive science, and cultural psychology (HIP13: 17). What made these frontiers exceptional was the fact that "researchers in other disciplines are joining forces with psychologists to forge new approaches to the study of psychological phenomena" (HIP13: 16). As for the main approaches/perspectives, the authors cautioned the students against privileging one over the other, because "these approaches need not be mutually exclusive; rather, they may focus on different aspects of the same complex phenomenon" (HIP10: 9). In the seventh edition, the authors called such a stance an "eclectic approach", which "uses a synthesis of several viewpoints explaining different psychological phenomena" (HIP7: 4).

The approaches/perspectives were written as a convergent and trimmed-down overview of the discipline. The neurobiological approach was discussed first, framing psychology into a physicalist worldview, in order to clearly separate psychological research from potential dualism: "Ultimately, it may be possible to specify the neurobiological mechanism underlying even the most complex human actions. However, a comprehensive neurobiological theory of behavior is at present only a remote possibility" (HIP7: 5). The same module exists in the sixth edition, but it mentioned "a comprehensive neurobiological theory of man" (HIP6: 6) instead.

In the tenth edition in 1990, the biological perspective was qualified more cautiously: "Because of the complexity of the brain, tremendous gaps exist in our knowledge of neural functioning. For this reason, a psychological conception of ourselves based solely on biology would be inadequate" (HIP10: 10). The authors employed the 'unfinished science' strategy, but did not wholly reject the idea of a complete reduction of the psychological to the neurobiological. This hedging completely disappeared from the eleventh edition, which presented the biological perspective in a much stronger light than the older editions. However, the later editions (HIP10, HIP11, HIP12, and HIP13) all had an explicit section right after the description of all the approaches/perspectives in which they elaborated on the question of the relationships between the various perspectives. The most contentious relationship was between the (neuro)biological perspective and all the others, considering they operated on different levels as explanations. The other four (behavioral, cognitive, phenomenological/humanistic, psychoanalytic) offered psychological explanations, while the biological one produced physiological and neurological explanations. The difference between them broached the question of reduction. And even though the authors themselves presented successful examples of reduction in the introduction and throughout the book, they did not subscribe to an all-out reductionist program for psychology: "[I]s psychology just something to do until the biologists figure everything out? The answer is no" (HIP13: 14; a variant of this module appears in HIP10: 14; HIP11: 12; HIP12: 15).[68]

The relationships between these perspectives were a source of structure for the exposition of psychology in the whole book. For the authors, the connections between perspectives could only be interpreted consistently if one presupposed that psychology was an unfinished science:

> The perspectives are competitive when they offer different explanations for the very same phenomenon. This kind of conflict will arise many times throughout this book. Such a conflict may indicate only that our knowledge of the relevant phenomenon is imperfect. As more is learned about the phenomenon, the views may become compatible with one another. An initial conflict among the views may thus be just another step in the ongoing process of scientific psychology. (HIP10: 15; variant of the fragment in HIP11: 13)

A consistent thread connecting the introductions in Hilgard's textbooks through the decades – whether they were organized around a historical overview or the subject matter – was the coverage of methods psychologists used in their research. To this group of modules I turn in the last part of my analysis.

3.2.3 Methods

The most stable part of the introductions through the decades of Hilgard's textbooks was the section giving an overview of research methods. The method modules were consistently present from the first editions, bringing "the methods of the experimental laboratory", "naturalistic observation", "case histories", and "test methods" (HIP1, HIP2, HIP3, and HIP4), with the addition of "the interview" in the second to fourth edition, and a separate discussions of "statistical methods" in the second edition that

[68] The complex relationship between the brain and the *psy* sciences is a large topic in historical research (for a general discussion see Vrecko, 2010; for an illuminating case study, see Weidman 1999). Since this complex topic is not in the focus of this chapter, I will not expand on it even though textbooks might be useful sources for investigating the boundary-work and especially scholarly identity construction in the space between brain and mind sciences.

warped into a whole new section of the introduction under the heading of measurement (HIP3 and HIP4). In the paragraph introducing the method section in the first edition, Hilgard commented that the “named systems” of theories have become less important for psychologists, and “attention has turned increasingly to efforts to improve the methods of psychology as science, to improve its operations” (HIP1: 568-569). And indeed, the view of psychology reconstructed from the introductions of his textbooks in this article corroborates exactly that.

In the first four editions, naturalistic observation was described as a method appropriate for “the early stages of science” when “it is necessary to explore the ground, to become familiar with the relationships that will later become the subject of more precise study” (HIP4: 11). Case histories were described as a type of “scientific biography”, which were “developed largely in connection with social work” (HIP3: 13).

The queen method was, of course, the experiment because “[w]e distinguish the methods of the experimental laboratory from naturalistic observation, case studies, test approaches by the *control of variables*” (HIP1: 571, emphasis in the original). The author described what were independent and dependent variables, noting that the “[c]ontrol of variables, as well as quantitative measurement, characterize true experimentation” (HIP1: 571). And even though the author took pride in listing the “precision instruments” one would have found in a psychological laboratory (very much “like other science laboratories”, Hilgard stressed), the true value of psychological experimentation was in the logic of its setup, not in the materiality of its instrumentation:

> The value of an experiment is not determined by the amount of apparatus used. Fundamentally, arrangements for experimentation are a matter of logic. Experiments have to be carefully designed for their results to be informative in relation to some hypothesis or theory. If the logic of experimentation requires precision apparatus, then such apparatus is used; if it does not, then good experimentation may be done with pencil and paper. (HIP1: 571)

All four of the first editions contained this module, cushioning it with the caveat that “for psychology to develop as a science, it is not essential that all its problems be brought into the laboratory”, especially considering that other sciences “such as geology and astronomy, are experimental only to a very limited extent” (HIP3: 11).

The mention of test methods in the first and second edition was expanded into a whole five page measurement section in the third and fourth edition, discussing experimental designs and “the use and interpretation of correlation coefficients” (HIP4: 25). Correlational methods were introduced as the next best thing when experimental controls were out of reach: “Although the experimental ideal is to have things so under control that one can specify the variables under study and produce changes in them as called for in the experimental design, there are circumstances when this is not possible” (HIP4: 25).

The fifth edition renamed the methods modules, giving the subheadings “methods of the experimental laboratory”, “field observations”, “survey methods”, “test method”, and “case histories and longitudinal studies”, keeping the section on measurement in psychology largely intact. The change from previous editions was mostly in order and subheadings, while the modules describing the particular methodologies stayed the same. Psychology was still introduced as a science of variables, for example: “The distinguishing characteristic of a laboratory is that it is a place where conditions can be carefully controlled and measurements taken, so that regular (“lawful”) *relationships among variables* can be discovered” (HIP5: 13, emphasis in the original). Experiments were contrasted to survey methods, with the example of Kinsey’s studies of sexuality (HIP5: 14).

The sixth, seventh, eighth, and ninth edition kept a similar structure – the Methods section bringing “experimental method”, “observational method”, “survey method”, “test method”, and “case histories”; followed by the measurement section. The modules bringing the listed methods largely stayed the same, in some cases illustrated by new examples. Thus, a referenced example for bringing the observational method to the laboratory was Masters and Johnsons study of sexual response (HIP6: 18). The same examples from sex research – Kinsey and Masters and Johnsons - appeared in all of the following editions in my analysis, testifying to how stable the method section across the editions really was.

In the tenth, eleventh, and twelfth editions, the authors grouped the methods under three large subheadings: “experimental method”, “correlational method”, and “observational method”; with the thirteenth edition keeping the same structure but renaming the large heading of Methods to “How Psychological Research is Done” (HIP13: 15), and relabeling the subheadings to “Experiment”, “Correlation”, and “Observation”. In these last editions in the 20th century, the modules introducing methods were updated from the previous versions, but they consistently introduced the same content. The novelty was that methods were explicitly connected to the previously introduced perspectives (what was called approaches in the older editions): “While some of the methods are used more by certain perspectives than others, each method can be used with each perspective,” the exception being phenomenological psychologists who “reject scientific methods entirely” (HIP10: 15). By the 1990’s, the dominance of quantitative psychology of variables - in the view of the textbooks authors - was so absolute that research traditions falling outside of the methodological standard were not even perceived as scientific.

The experimental method was introduced in largely the same way as before, with paragraphs on the control of the relationships between variables (HIP10: 16), and the same examples of experiments on the effects of marijuana consumption that appeared in the previous editions. The correlational method subheading brought the content which used to be in the separate measurement section in the older editions, while the observational method section kept all the rest: Direct observation, survey method, and case histories. All in all, what happened in the tenth, eleventh, and twelfth edition was a merger of the methods with the measurement section. A new appendix appeared at the end of the book's eleventh edition, which brought in even more details about the statistics used by psychologists, covering descriptive statistics and statistical inference.

Lastly, the opening of the method section in the twelfth and thirteenth edition received a new short module about generating hypothesis, where the authors instructed the student that generating new hypotheses came from two possible sources: Either from being "an astute observer of naturally occurring situations" or "being a scholar of the relevant scientific literature" (HIP12: 17). The suggestion to use the vast literature for hypothesis generation is very much in line with the continuously present awareness of the size and constant expansion of psychological literature.[69]

The continuity of the methods introductions in the textbooks is an expected feature of post-WWII psychology. The stabilization of the methodological core of the discipline, especially in the United States, is a common observation among historians of psychology. As Jill Morawski (2005, p. 99) describes it:

> While laboratory experimentation became the principal method of inquiry, other methods (notably survey, correlation, and individual difference studies along with observational techniques) did not disappear. However, these other methods took second stage in textbook descriptions of psychological research methods and increasingly came to be evaluated in terms of their degree of adherence to some of the core ideals of laboratory experiments.

Historical studies taking this view on the development of late-modern psychology abound – from Kurt Danziger (1997, p. 173) calling the new language of variables the *lingua franca* of psychology, to specific studies of its institutionalization in psychology through textbooks (Winston, 1990; Winston and Blais, 1996) and journals (Danziger and Dzinas, 1997). Hilgard's textbooks in the period from 1950 to 2000 corroborate the view that psychology's research methods were highly regularized, institutionalized, and delimited into a rigid epistemological pecking order – laboratory experiment with its control of variables on top, correlation methods of analyzing naturally occurring variability just below, and everything else in the service of those two.

---

[69] Mentions of the constantly expanding scientific literature of psychology were a staple of the textbooks prefaces. For example: "With the growth of the psychological profession and the increasing financial support for research, the amount of new research bulks large indeed, with over 10,000 new titles reported annually in *Psychological Abstracts*. The problem of selecting from this vast body of material becomes a staggering one" (HIP3: preface, para. 2).

3.3 The what for?

I will recapitulate some of the main points in my chronological analysis of the introduction rhetoric of psychology in Hilgard's textbooks and then draw some more general conclusions.

The unfinished science in a changing society. The way that Hilgard and his co-authors opened up their introduction changed through time. From the first two editions that were student-friendly with the story about a science arising from folk psychological concepts, the editions in the 1960s and the 1970s gave way to the more academic exposition on the reflexive relationship between how psychological knowledge was produced and what it was. From 1975 onwards, the psychologized society reigned supreme. There was no need for justificatory academic foundations, because psychology, as a science and profession, was present in so many facets of society. The common thread through these different views on psychology was that the discipline– whether it arose from folk psychology, was an inevitable consequence of the human mind, or a practical necessity – was unfinished. Why did the rhetorical strategies for the opening paragraphs change? I would speculate that this has to do with the institutional history of the discipline in the United States, for example, psychology becoming (in sheer proportion of employment of psychologists, but also in its content) less of an academic field and more of a profession (Green, 2015, p. 210).

The trimmed discipline overviews. The point about psychology being unfinished might have been subtle in the pitch, but the way the body of the introduction changed from historical overviews to the approaches/perspectives and professional specializations shows it in full view. In the editions up to the 1970s, psychology was presented as a hard-won triumph of empirical science. However, the success story was qualified – the state of theories and explanations psychologists came up with made the authors cautious. Maybe psychology was not a philosophical discipline anymore, but in their view, it wasn't a mature science either. The promise of a resolution lay in the rigorous application of methods. The convergent narrative of approaches/perspectives in the 1980s and 1990s was the culmination of the view Hilgard expressed about S-R psychology in the 1960s: Entrenched theoretical positions brought metaphysical baggage, and good scientists should avoid metaphysics. The only grounding a researcher in psychology needed was the abstracted language of operationism and variables. That language made the growingly sophisticated statistics and methodologies universally intelligible to psychologists. Approaches and perspectives represented stances that could be slotted into this uniform system of thinking, according to the preference of the investigator. The fact that the *lingua franca* of variables excluded incompatible psychological systems of thought was not an issue, because it rendered them invisible. It pushed them outside of the scientific consensus, as in the mentioned example of what the textbook authors called phenomenological psychology.

The stability of methods. For mainstream psychologists in the second part of the 20th century, the disciplinary consensus about psychology was an agreement on methods. Zooming out of the constrained perspective constructed from textbooks in this chapter, there is empirical proof that psychologists publishing in English-language psychology regimented their area of investigation according to strict methodological rules. The scientometric study in Chapter 4, based on the term-mining of hundreds of thousands articles published in psychology's varied research lines, shows that psychology's literature in the second half of the 20th century exhibited a stable structure, and that the structure represents the methodological language of psychology. In other words, in our analysis of the literature, we find the same methodological core that the introductions to Hilgard's textbooks kept as a stable thread through fifty years.

This is not to say that there was no methodological innovation during the second part of the twentieth century – it's easy to come up with two examples as new approaches that were developed and spread through quantitative psychological proper: Structural equation modelling and Bayesian statistics. But those innovations were extensions of the methodological core that was agreed upon – they were an increase in sophistication, not a reform in research design.[70]

3.4 Conclusion

The crucial feature that can be retrieved from these textbook views of psychology was the psychologists' ability to translate all the varied theoretical and metaphysical positions into a singular language of methods. Paraphrasing Danziger: Instead of talking about personality, psychologists would talk about personality variables; instead of intelligence, one had intelligence tests. Theories were not reduced and constrained by accident; it was the design feature of the psychologists' view. The commitment to an internally consistent system of experimental design, variables, and statistical correlation of late 20th century psychologists can be interpreted as what Peter Galison (1987, p. 246) called "long-term constraint". In Galison's appropriation of Braudel's view of historical time, "long-, medium-, and short term constraints […] articulate the ways in which prior experimental and theoretical beliefs narrow alternatives of what the experimentalist takes to be reasonable beliefs and actions."

What are the historiographical implications of this study? I hope to have argued convincingly that the ensemble of rhetorical strategies used to introduce psychology to generations of students was exactly the "narrowing of alternatives of reasonable belief and action" for psychologists in the second part of the 20th century. That

[70] An excellent example of this increase of sophistication is the recent history of construct validity theory (see Slaney, 2017). Construct validity theory was not overturned or fundamentally altered since its institution in the 1950s, it was extended and increased in sophistication.

narrowing is so evidently clear and straightforwardly reconstructed from textbooks precisely because these books are reified personal, yet expert, views about the discipline that are directed at laymen. In other words, the very distortion in perspective is an inroad for historical analysis and for drawing substantive conclusions.

The substantive conclusions are as follows. In the view reconstructed from Hilgard's textbook, the psychologists' rules for thinking scientifically were "acquired without notice in learning the theory of a given epoch" (Galison, 1987, p. 246). But unlike Galison's science of high-particle physics and the interactions between groups of theoreticians and experimentalists that underwrote it, psychologists did not receive their methodological constraint through the medium of the dominant theory. They received it through a strict epistemological attitude as to what counts as science. The greatest achievement of late-20th century American psychology was not in its theories or concepts, but in the regimented thinking psychologists imposed upon themselves. Like the Christian monks of Late Antiquity who invented introspection to produce psychological knowledge by systematic imposition on the way they thought (Graiver, 2017), psychologists of the 20th century have succeeded in a similar thing. The biggest difference being in the cultural framework – for psychologists it wasn't the theological context which focused investigation on the subjectivity of a unique soul endowed by God, but the subjectivity of an evolved mind being explained by natural laws.

Following my analysis based on textbooks, the replication crisis in 21st century psychology can be perceived in two different ways. The first reading is that the replication crisis is a natural outcome of the epistemological program of psychology, and that the program's strongest advocates are trying to reign in the throngs of psychologists who follow the program inconsistently. The upstart methodological elite is cleaning up the ranks and trying to enforce stronger norms for valid research. The second reading, the more radical one, does not buy into the discipline's history of progress. Only on the surface can the replication crisis be read as reformers strictly reinforcing the rules of good research that were established decades ago. In this view, the replication debates are the most recent crisis of a metaphysical program investigating human psychology that has realized it has no tools to fix itself, because methods alone do not solve methodological problems.

Chapter 4. Framing psychology as a discipline (1950 – 1999)

"Excitement and pink lemonade." That's how Lee Cronbach described the state of psychology in his presidential address, on September 2, 1957, to the assembled members of the American Psychological Association:

> No man can be acquainted with all of psychology today, as our convention program proves. The scene resembles that of a circus, but a circus grander and more bustling than any Barnum ever envisioned – a veritable week-long diet of excitement and pink lemonade. Three days of smartly paced performance are required just to display the new tricks the animal trainers have taught their charges. We admire the agile paper-readers swinging high above us in the theoretical blue, saved from disaster by only a few gossamer threads of fact, and we gasp as one symposiast thrusts his head bravely between another's sharp toothed jaws. This 18-ring display of energies and talents gives plentiful evidence that psychology is going places. But whither? (p. 671).

This was the key question: whither? In other words, where was psychology going? Or, more generally: What was it that was going? And thus, similarly: Where has 'it' been? Cronbach's metaphorical circus is a common *topos* for many psychologists, but usually under a more down-to-earth name: Psychology's crisis of disunity. Much has been written about it since the 1950s. And not only by psychologists - just looking at the history of the naming conventions trying to delimit psychology as a science is indicative of how interesting this question was and is for historians and sociologists.[71]

Cronbach, in 1957 and in 1975, described what he called the two disciplines of scientific psychology. The schism he identified was between "correlational" and "experimental" psychology. On the correlational side, psychologists looked at the existing differences

---

[71] For a sample of the many treatments of disunity by psychologists, see Staats (1983, 1991), Koch, (1993), Koch and Leary (1992), Green (1992a, 2015), Stam (2004), Sternberg, (2005), Henriques (2003); see a special section of *Theory & Psychology* dedicated to Henriques' ToK project (Henriques, 2008). The question of unification has often been labeled as a crisis (e.g. Goertzen, 2008). For more on the intellectual and social histories of these umbrella terms naming psychology, see the *Introduction* in Erickson et al. (2013) on moral/social/behavioral sciences; Pooley and Solovey (2010) and Pooley (2016) on behavioral/social sciences; and for the inter- and multidisciplinary history of the cognitive sciences, see Cohen-Cole (2007). For a Foucauldian conception of 'psy sciences' see Nikolas Rose (1990, 1998). These labels are not only different names for psychology - they also act as conceptual workhorses for wider boundary-making (Gieryn, 1999) of psychology's association with different social sciences, philosophies, natural sciences, and medicine.

between subjects and devised statistical procedures to analyze them. The data on which their science was built were the many variables on which people could differ – everything from intelligence to various measures of personality. On the experimental side, psychologists designed experiments in which they tried to keep all things constant and isolate the influence of a particular experimental intervention. The data about how that treatment affected the subjects - and how varying the treatment the experimenter could get different results - was the bedrock of their approach. These two broadly sketched lines of research, according to Cronbach, were methodological traditions with their own histories, communities, and rules.

Was Cronbach's description warranted? Considering the many pitched debates among psychologists about the nature and boundaries of their discipline throughout the twentieth century, it sounds plausible. Cronbach felt it was one of the most important questions to tackle at the time – in his previously quoted presidential address, he wrote that the different methodological traditions limited psychology, and that investigators should dedicate themselves to "scientific psychology as a whole" (1957, p. 671). His later evaluations were less optimistic (Cronbach, 1975). Even today, some methodologists agree with Cronbach's description of the state of research in the field (Borsboom, Kievit, Cervone, & Hood, 2009).

Considering the popularity of the unity/disunity debates among psychologists and historians, we approach the question of disciplinary formation by reframing it into a history of methods. Such a move is not without precedent – if we take a look at relevant historiography much work has been done on the history of psychology's methods in the 20th century by Andrew Winston (1988; 1990; 2005; MacMartin & Winston, 2000), Kurt Danziger (1985; 1990; 1996; Danziger & Dzinas; 1997), and Henderikus Stam (1992; 2000; 2004). They have addressed the role of the methodological meta-language, particular research designs, and statistics used in psychological research. Their work brings to light the development of methodological uniformity in psychological research in the period, but they do not frame their analysis along the lines of Cronbach's correlational/experimental distinction. If there is a kind of methodological uniformity developing in the late 20th century psychology, how does this uniformity fit Cronbach's two streams of scientific psychology? And even more fundamentally, how does the historians' idea that research methodologies were converging fit with Cronbach's (and other psychologists') narrative of disunity? We aim to explore exactly that through a large-scale analysis of the content of psychological journals.

Cronbach's two disciplines of scientific psychology and the work of Danziger, Winston, and Stam are our starting point for a *bibliometric* analysis. Historians have identified the philosophical and social forces underscoring the methodological imperative internalized by psychologists in the 20th century. In the same period, the research output of psychologists experienced a staggering growth, as is the case for most of science (see De Solla Price, 1963/1986; also see *Figure 4.1* in this chapter). Psychology, the "traditional history" of the discipline in the 20th century would tell us

(Walsh, Teo, & Baydala, 2014, p. xiii), also went through important changes – the cognitive revolution, the rise of evidence-based therapies and various professional psychologies, the advent of neuroscience; just to name a few. If we take a bird's-eye view of psychological research, made possible by new ways of analyzing large amounts of data, can we identify a) the growth of the literature b) the fundamental changes in the content of the science c) the methodological traditions akin to the ones Cronbach talks about?

We aim to show that even though the growth of the literature was massive, the *fundamental changes* in the content of psychological research were not structural. The structure of the field remained the same, and at the center was a methodological core. As for the content of psychological research (whether we call it theories, paradigms, or psychological knowledge), we can hardly talk about psychology expanding; it is more appropriate to talk about facts accumulating;[72] facts which are generated and justified within a closed system of supposed methodological uniformity. The scientific edifice in the 1990s has become larger, but it is structurally very much alike to what Cronbach saw in the 1950s. Our study is a first of its kind in trying to document the supposed disunity of late 20th century psychology through empirical methods of analyzing the scientific literature on a massive scale.

4.1 Method

Articles published in *History of Psychology* (both the journal and the historical discipline) usually do not have a method section.[73] We have decided to include one despite it being uncommon, agreeing with Green, Feinerer, and Burman (2015a, p. 17) that "it seemed advisable [to include a Method section] because we used a set of technical procedures that are unfamiliar to most historians…[and] an explicit "Method" section seemed to be the most efficient way to convey this information." Here we will introduce the dataset we are working with, the rationale behind turning to digital history, and the computational tools we use in our analysis.

Our approach is based on data-mining terms from scientific journals. We use the VOSviewer software developed for scientific literature analysis by scientometricians (Van Eck & Waltman, 2010, 2014c; http://www.vosviewer.com). Instead of coaxing out future research fronts, we turn our gaze backwards and use the same tools to

[72] A good metaphor for this kind of haphazard accumulation of facts is an "exploding confetti factory" which Ruud Abma (2013, p. 115) discusses in his book on the fraudster Diederik Stapel. The knowledge explosion was originally described as a confetti factory in a book review by Barclay (1973). Collections of facts generated by psychological research cannot be called paradigms or theoretical systems. They are collections bounded by certain methodological and institutional traditions. For more on what it means to generate and constitute facts in psychology, see the work of Mary Smyth (2001a; 2001b; 2004) and the critical synthesis on psychology's textbook fact-making in Chapter 2.

[73] This chapter was published as an article in the journal *History of Psychology*.

reconstruct research fields in the past. We take a sample of 676,393[74] articles published in journals indexed in PsycINFO[75] from 1950 to 1999, and conduct an analysis of the relevant terms they use in their abstracts and titles. These terms are visualized in two-dimensional co-occurrence maps of the discipline in the following way: The larger the number of abstracts/titles which contain the same two terms together, the closer those terms will appear in the map. In this way, the abstracts/titles are used to generate terms, then the co-occurrences of these terms structure them into a map. We see the co-occurrence maps as a proxy for the discipline of psychology in the period when the articles in our corpus were published.

4.1.1 Arguments for digital history

Why use term-mining? When Green, Feinerer, & Burman (2015a) used "distant reading" to analyze the trends in *Psychological Review* from 1894 to 1908 they made the argument that "there is far too much source material to be handled by the means that historians traditionally use." They describe the problem as follows: "To capture it all, the individual historian needs a way of handling, organizing, and manipulating this large mass of historical material without having to individually read, interpret, and situate each of these thousands of items" (p.16). The problem has become exacerbated since the 1950s when our analysis starts, because the production of literature, in absolute numbers, doubled with each decade of the second half of the twentieth century. Given the literature explosion, using digital tools to try and analyze disciplinary formation is not just a novel tool, but a crucial one.

However, literature size is not reason enough to turn to digital humanities. The number of texts was almost always too big for comprehensive overviews – the meaningful synthesis of such unsurveyable amounts of information is the bread and butter of historians of science. A more compelling reason for taking the digital approach to historical analysis of disciplinary formation is that it allows us to take a perspective that is not based on prominent authors, their publications, and the fields with which they were associated. There is a certain democracy of large numbers involved in taking the term-mining perspective, where the terms that dominate texts

---

[74] The present study is the first of its kind considering the scope: We cover as much psychology as possible to extract large-scale historical trends from the literature. Pioneering work on applying digital analysis to history of psychology has been done by Christopher Green's group of historians at York University (Green, 2016; Green, Feinerer, & Burman, 2013, 2014, 2015a, 2015b; Pettit, Serykh, & Green, 2015; Burman, Green, & Shanker, 2015; Young & Green, 2013; Green & Feinerer, 2015), but on a smaller scale and in a different time period.

[75] An interesting approach to discourse-vocabulary-discipline investigation is the work of John Benjafield (2012, 2013). He also uses PsycINFO and tries to historically investigate psychology through the terms used by psychologists but his approach is quite different from ours. Closer to our work is that of Burman, Green, and Shanker (2015), who investigate self-regulation using PsycINFO's controlled vocabulary. The main difference being that our vocabulary is text-mined, not controlled.

frame our view - not analytical categories like individual or institutional reputation. The *Big Names* still exert their influence over historical trajectories of terms by virtue of their importance, but by taking the term-mining road, they are not the point from which we start as historians.

Considering the above, we can take a very minimal definition of what psychology was in the period 1950-1999: Whatever the academics and practitioners in the examined period called by that name. More than a thousand journals are included in the analysis.[76] We were guided by maximum inclusivity - to include as many voices published in academic journals on psychological topics, and then analyze them *en masse*. The idea behind the approach is to control for essentialist perspectives on what the core of the discipline was. We subscribe to a rudimentary empiricism of the digital age: We do not focus on experimental psychology, or behaviorism, or the cognitive sciences, or the various applied psychologies - we include as many data points covering all of them and more, and then analyze the patterns that arise out of the data.[77]

As far as our philosophical position toward digital humanities goes - what are the patterns we are analyzing and are they good proxies for disciplinary formation? In that, we are dynamic nominalists (Hacking, 2006).[78] The mass of dots in the maps presented here are constellations of terms that contain echoes of intellectual traditions of psychologists. We are picking up bits and pieces – several hundreds of thousands - of the writing found in journals and trying to fashion an abstracted story of what psychology was in the second half of the twentieth century.

4.1.2 Data

The first step in our study was the construction of a representative dataset of the psychological literature. To construct this dataset, we used PsycINFO, a bibliographic

[76] For a full list of journals and the number of publications in the data set see https://figshare.com/account/projects/16467/articles/4232273

[77] A historian of psychology might read our two reasons for using digital methods as unqualified endorsements of the *New History of Psychology* (Furumoto, 1989; and a re-evaluation in Lovett, 2006). Although we share some views with Laurel Furumoto, we do not see various approaches to history of psychology as incompatible. Our approach is yet another contribution to history of psychology as a “divided discipline” (Weidman, 2016b, p. 252), adding another tool to the “historian’s toolbox” (Green, 2016; p. 218).

[78] Digital methods quite literally make us dynamic nominalists because we deal with ephemeral names of things that appear in the summaries and titles of published literature. The psychologists’ research “objects”, subjects who acted as sources of data, researchers with their idiosyncrasies in their institutional and social contexts; all remain an elusive background to the thousands of names of things we will showcase in the analysis. Our approach does not invalidate all these objects and subjects behind the literature as crucial elements of causal explanations of historical development. On the contrary, the data-mined perspective offers a framework for explicating them through conventional historiography.

database containing the meta-data of more than four million articles from the field of psychology and related disciplines. The dataset we extracted from PsycINFO includes most indexed articles of the document type 'journal article' published between 1950 and 1999. The number of articles in the dataset is 676,393. *Figure 4.1* provides a breakdown of the number of articles for each of the five decades. It can be seen that over time there has been a rapid increase in the number of articles. The articles appeared in 1,269 different journals. *Table 4.1* provides an overview of the 20 journals with the largest number of articles in the dataset (for the full list of included journals, see footnote 76).

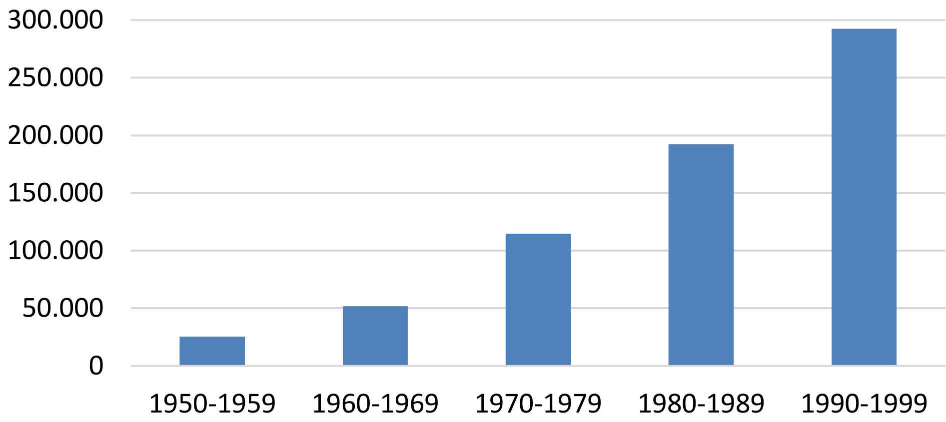


Figure 4.1. Number of articles per decade in the dataset of psychology articles

Ideally, we would have preferred to work with the full text of articles included in our analysis. Unfortunately, however, we could only use article titles and abstracts because of limited database access. Our data was manually retrieved from PsycINFO with a lot of effort and man-hours, focusing on retrieving article-level metadata from the database (e.g. title, abstract, and years of publication, among other things). Retrieving full texts on the same scale is currently impossible. The newly founded PsycINFO Data Solutions service (http://www.apa.org/pubs/psycinfodatasolutions/) is supposed to provide the kind of access we would need for our research, but the level of access for large-scale discipline-wide literature analysis is still well beyond the capacity of this service (PsycINFO, personal communication, November 8, 2016).

Table 4.1. Top 20 journals with the largest number of articles in the dataset of psychology articles.

Journal	**No. of pub.**
Psychological Reports	17,761
Perceptual and Motor Skills	16,291
Physiology & Behavior	9,736
Brain Research	7,554
Pharmacology, Biochemistry and Behavior	7,159
Journal of Personality and Social Psychology	6,899
Journal of Clinical Psychology	6,786
American Psychologist	6,400
Psychopharmacology	5,997
Animal Behaviour	5,972
Journal of Experimental Psychology	5,636
Journal of Comparative and Physiological Psychology	5,484
Child Development	5,480
The Journal of Social Psychology	5,000
Educational and Psychological Measurement	4,935
Perception & Psychophysics	4,738
Journal of Applied Psychology	4,726
Journal of Nervous and Mental Disease	4,692
Journal of Consulting and Clinical Psychology	4,368
Vision Research	4,331

4.1.3 Term identification

After collecting the titles and abstracts of the articles in our analysis, we tried to identify the terms that best capture the intellectual structure of the psychological literature. This was done by following the automatic term identification approach developed by Van Eck & Waltman (2011). This approach automatically identifies relevant terms in the titles and abstracts of the articles in our dataset. More specifically, using natural language processing techniques, we first identified noun phrases in the titles and abstracts of the 676,393 articles in our dataset. A noun phrase was defined as a sequence of words such that the last word in the sequence is a noun and each other word is either a noun or an adjective. We then converted plural noun phrases into singular ones. In this way, noun phrases like "symptom" and "symptoms" were unified. Only noun phrases occurring in the titles and abstracts of at least 300 articles were taken into consideration. Noun phrases occurring in fewer than 300 articles were excluded in order to keep the number of noun phrases included in the analysis manageable.

In visual analyses like the one presented in this paper, it is typically not useful to include more than a few thousand noun phrases. The requirement that a noun phrase needs to occur in 300 articles resulted in a set of 4,913 noun phrases. Of these noun phrases, the 3,000 noun phrases that seemed most relevant were selected. The selection was made based on relevance scores calculated using a computer algorithm (Van Eck & Waltman, 2011). These relevance scores were used to distinguish general noun phrases with a broad meaning (e.g., "method", "result", and "conclusion") from more specific noun phrases (e.g., "depression", "memory function", and "posttraumatic stress disorder"). The latter noun phrases tend to be the more interesting ones, and these noun phrases were therefore selected. In our experience, selecting about 60% of the noun phrases typically works reasonably well. Although our algorithmic approach to select noun phrases does not always give optimal results (some relevant noun phrases may be excluded from the selection and some non-relevant ones may be included), the selected noun phrases generally can be regarded as important and relevant terms in the field of psychology.

4.1.4 Term maps

Based on our set of 3,000 important and relevant terms, the next step was the construction of the so-called term maps. A term map is a two-dimensional visualization in which the terms are located in such a way that the distance between any two terms reflects the relatedness of the terms as accurately as possible. In general, the larger the number of co-occurrences of two terms, the smaller the distance between them. In this way, a term map provides a visual overview of important topics discussed in the literature and how these topics relate to each other. The larger the number of articles in which a term occurs (in the title or abstract), the more prominently the term is displayed in a term map. Frequently occurring terms are for instance presented using a larger font size.

In total, we constructed six term maps. To get a general impression of the subdivision of the field of psychology into topics, and how these topics relate to each other, we constructed an overall term map based on all articles in our data set. To identify changes over time in the topical focus of the field, we also constructed term maps based on the articles published in each of the five decades covered by our data set, that is, 1950–1959, 1960–1969, 1970–1979, 1980–1989, and 1990–1999.

The six term maps were created using a software tool called VOSviewer (Van Eck & Waltman, 2010, 2014c; http://www.vosviewer.com), in which VOS stands for visualization of similarities. To construct a term map, we determined for each pair of terms the co-occurrence frequency. The co-occurrence frequency of two terms is obtained by counting the number of articles in the relevant time period in which the two terms both occur (in the title or abstract). We then used the co-occurrence frequencies of the terms as input for the VOSviewer software. Based on these frequencies, the VOSviewer software constructed a term map in which the distance between any pair of terms provides an approximate indication of the relatedness of the terms as measured by co-occurrences. Each term in a term map also has a color. Colors are used to indicate the grouping or clustering of terms into topics. Terms that have the same color belong to the same cluster and tend to be more closely related than terms having different colors. In other words, terms that have the same color tend to co-occur with each other more frequently than terms having different colors.

To obtain the layout and the clustering of the terms in a term map, the VOSviewer software uses a mapping technique and a clustering technique. These techniques jointly provide a unified framework for mapping and clustering (Waltman, Van Eck, & Noyons, 2010). The mapping technique determines the layout of the terms in a term map (i.e., the locations of the terms in the map), while the clustering technique produces a clustering of the terms in a term map by assigning frequently co-occurring terms to the same cluster. In the example presented in *Figure 4.2*, the layout of the terms was determined by the mapping technique while the clustering of the terms, indicated by colors, was produced by the clustering technique. The techniques were applied independently for each of the six term maps based on co-occurrence frequencies obtained for the relevant time period. Each map therefore has its own layout and clustering.

The mapping technique used by the VOSviewer software is called VOS. This technique is closely related to the technique of multidimensional scaling (e.g., Borg & Groenen 2005). We refer to Van Eck, Waltman, Dekker, & Van den Berg (2010) for a discussion of the advantages of the VOS mapping technique over approaches based on multidimensional scaling. The VOS mapping technique has the so-called attraction and repulsion parameters that allow for some degree of customization in the way terms are positioned in a term map. We used a value of 1 for the attraction parameter and a value of 0 for the repulsion parameter. These values yielded the most satisfactory layouts.

The clustering technique used by the VOSviewer software is closely related to modularity-based clustering (Newman, 2004; Newman & Girvan, 2004). For a detailed discussion of the clustering technique, we refer to Waltman et al. (2010). The clustering technique has a resolution parameter that determines the level of granularity of the clustering that is obtained. We used the default value of 1 for this parameter.

Term maps are invariant to rotation and reflection because these operations do not affect the distances between the terms in a term map. In order to facilitate easy comparison of the term maps obtained for the overall period and for the different decades, the locations of the terms in the various maps were aligned as much as possible. This was done using a technique known as Procrustes analysis (Borg & Groenen, 2005). Using this technique, the term maps were rotated and reflected in such a way that terms are located consistently in the different maps as much as possible.

In addition to visualizations in which terms are colored based on the cluster to which they belong, we also used the VOSviewer software to create so-called density and overlay visualizations. In a density visualization, colors are used to indicate the density of terms in the different areas of a term map. In an overlay visualization, colors are used to display the topical focus of the articles published in different sets of journals. Since the overlay visualizations that we use in this paper indicate frequencies of terms occurring in certain journals, we call them journal projections – they project specific journals onto the terms in a term map.

4.2 Psychology in term maps (1950-1999)

Our discussion of the term maps is structured in the following way. We first describe a basic pattern that arose out of the term maps. This pattern can be seen in each decade under investigation. We look into its most prominent characteristic - the structure of the two big superclusters - and then continue in analyzing these two superclusters with different sets of map overlays in the each of the decades.[79]

[79] The decade maps can be accessed directly and explored in detail online. This is recommended for a better understanding and using the full functionality of VOSviewer. Exploring each map with the zoom function and the ability to look at smaller groups of terms allows for easier and more intuitive inspection of the visualizations. The decade maps can be found here: 1950s: https://goo.gl/kqD8Ww 1960s: https://goo.gl/R3tKh4 1970s: https://goo.gl/yKNSgc 1980s: https://goo.gl/5TidHO 1990s: https://goo.gl/53SXFA. The overview map (1950-1999) can be found here: https://goo.gl/vPSAHC. For extra instructions on using the maps, see https://figshare.com/articles/Framing_Psychology_Map_Opening_Instructions/4043772

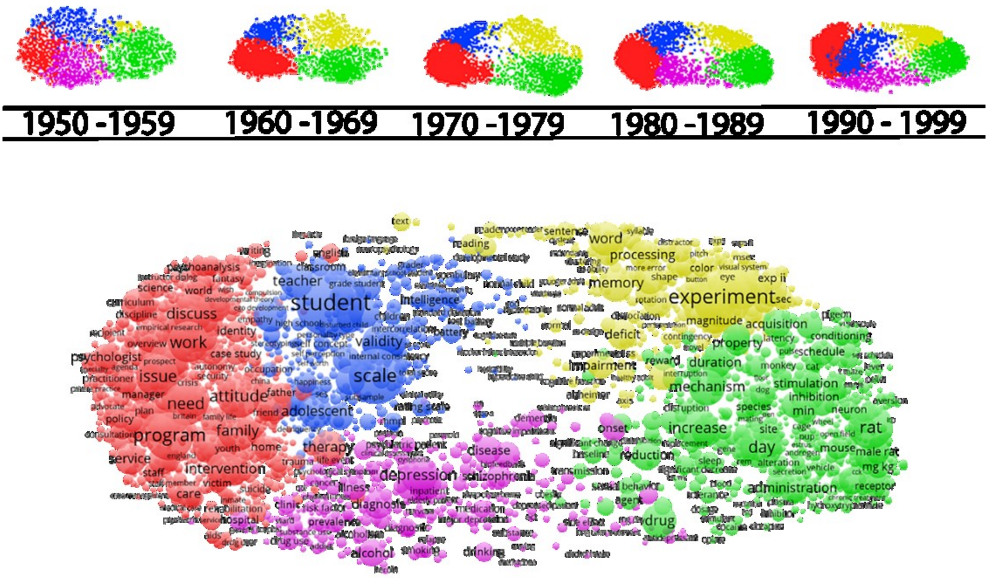


Figure 4.2. VOSviewer map with default clustering resolution (1.0)

4.2.1 Basic structure of psychology's term maps

The term maps based on our selection of journals exhibit a robust overall structure. We will explain it for the map of the whole period (*Figure 4.2* and its density visualization, which is represented by *Figure 4.3*). Each figure in the article will have the same layout. The terms in these maps were mined from all articles from 1950 to 1999 in our dataset. The bottom large map in the figure represents the whole period under analysis - from 1950 to 1999. The tableau on top of the map is a schematic representation of the occurrence of terms in each decade. In this way, each figure represents the whole period (the big map) and the schematic breakdown per decade (top band of schematic representations of the map).

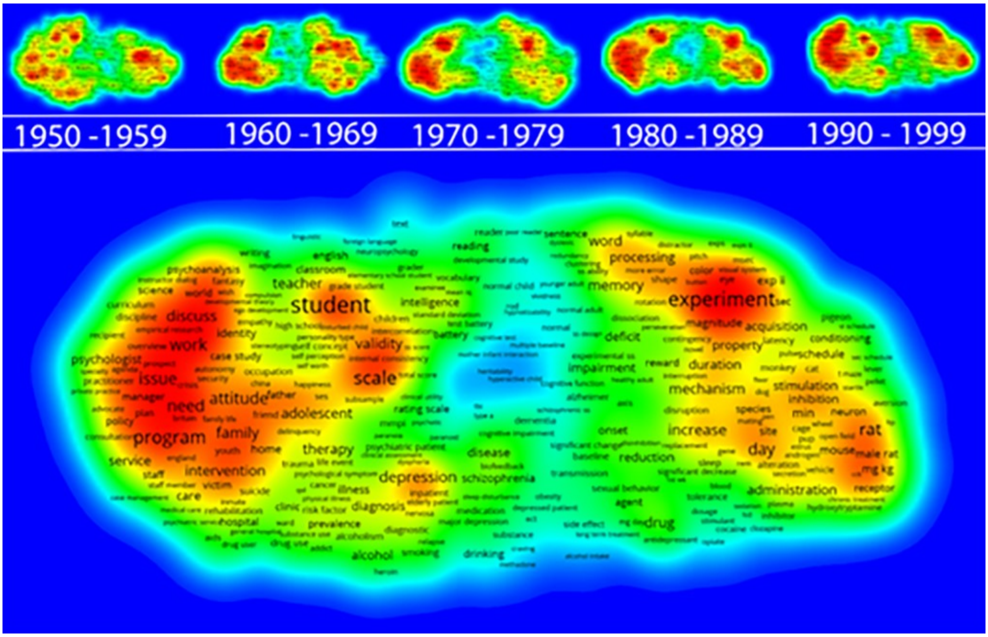


Figure 4.3. Density visualizations of VOSviewer maps

To begin with, we interpret the most robust feature of the maps. By robust, we mean the one that each of the maps exhibits. The field as a whole is structured into two superclusters - in *Figure 4.2* this can't be seen at first glance because we are using the clustering algorithm of VOSviewer. The interpretation of these clusters is slightly ambiguous, especially when we try to do it for different decades e.g. are the boundaries of the yellow clusters stable, and why does the purple cluster disappear in the 1960s and the 1970s?

But if we look at the density visualization in *Figure 4.3*, we see the pattern that the terms exhibit when they form more dense areas of the map. The 'warmer' the section, the more terms group in that area.[80] In the density visualization we can see that the terms group themselves in the western area and an eastern area, with an area of lower density in between. Here, the take home message is about the superstructure:

[80] Each point in the density visualization has a color that depends on the density of terms at that point. The larger the number of terms in the neighborhood of a point and the higher the weights of the neighboring terms, the closer the color of the point is to red. The kernel width (how far the 'redness' in the density visualization extends from a group of terms) is one of the manipulable parameters in making VOSviewer density visualizations. In this case, we have made it lower than the default to sharpen the distinctions between the smaller term clusters.

1. The map is divided into a western and eastern collection of denser clusters of terms. We call these denser clusters the eastern supercluster and the western supercluster. Supercluster is not a mathematical term; it just means that it encompasses different hotspots of terms that are close to each other but divided from the other half by the area of lower density in the middle of the map. Superclusters can be visually identified through the density visualization, or a low clustering resolution (the default in *Figure 4.2* being 1.0; if we lowered the resolution to 0.5,[81] it would clearly show the two superclusters). As the interpretation of these superclusters will show, they correspond to the various kinds of experimental psychology in the eastern supercluster and all the 'other' kinds of psychology in the western supercluster.[82]

2. In the center of the map, northern and southern bridges of lower density connect the two superclusters. For ease of identification, these are some of the terms in the northern bridge: "reading", "reader", "developmental study", "vocabulary". These appear in the southern bridge: "disease", "schizophrenia", "medication", "drinking".

Depending on the way terms were generated, the different values of the relevancy algorithm and occurrence thresholds, and the decade from which the titles and abstracts were selected, we get different terms in the maps. Thus, the structure of the terms' layout can be influenced by the parameters set for mapping and clustering. We varied all of these parameters and different approaches to the data, but all of them keep the superstructure of two distinct superclusters of terms: One on the west and one on the east, with a lower term density chasm between them. This consistent pattern allows us confidently to proclaim that we have identified a robust configuration that should be analyzed as a representation of the structure of psychology, at least as the underlying structure of terms appearing in titles and abstracts of psychological literature in English from 1950 to 1999. The caveat is that this analysis covers the psychology that is published in journals and somehow related to the mainstream traditions - and complementary traditions that flow alongside the mainstream - in English-language psychology. We include a small number of non-

---

[81] A map with a lowered clustering resolution can be found here: https://figshare.com/account/projects/16467/articles/4015695. You can manipulate the representations produced by different clustering resolutions yourself by opening the maps as linked in footnote 79. The clustering parameters in VOSviewer can be accessed in the left-hand tab under the label 'Analysis'.

[82] The interpretation of the meaning of the two superclusters will take the rest of this article. 'Experimental' and 'other' are placeholders for the discussion to come, aimed to help orient the reader. By placeholder, we mean a sort of promissory note (Manicas, 2006, p. 88): "a quasi or suggestive explanation" which still demands a further explanation of the underlying mechanisms producing it. By the end of the article, we will fill in the promissory note with an analysis of methodologies in psychology as the explanation of the stable structure.

English language journals (e.g. French, German, Spanish, Italian) that are indexed with translated titles/abstracts, but their number is too small to draw reasonable conclusions about psychologies in those languages. [83]

An important note is in order on the way visualizations are generated from our dataset. The York group of digital historians designates the visualization parameters under the researcher's control as less "objective" (see Green, Feinrer, & Burman, 2015a; p. 18-19) when they use Gephi (another software used for network and graph visualizations, www.gephi.org). In this, Gephi and VOSviewer are comparable, but we wouldn't agree that it makes the method less objective. It just makes it evident that we need to be aware of all the degrees of freedom we take in generating the maps, and being open about the choices we make, especially if they could have been different. Would taking different steps generate a different analysis? As far as we can tell: No. But this is an open question leading to a debate about large historical trends in literature. Differing interpretations are not only welcome, they are also necessary for digital humanities to become a robust and useful perspective among historians of psychology. Being able to vary these parameters would be less objective only if the algorithms and datasets we use are black boxes. If they are openly debated and examined, they represent the crux of digital historical scholarship. In our opinion, varying parameters is what makes tools like this interesting for generating historical interpretations. The distinction objective/subjective applied to methods is too unstable and might hide more than it explains. Because of this, we would strongly discourage the reader from viewing our approach as objective *or* subjective. We would rather call it data-driven and interpretive.[84] In using digital tools in history, we agree with Green's perspective that our aim is a well-designed visualization generating

[83] For a true international account of large-scale literature trends in psychology, the analysis cannot be done just in English despite its prestige and the amount of literature published in the language. Comparative studies should be conducted in other languages, or even better, in multiple languages at the same time. Comparative implies parallel development, while the case is probably extensive cross-pollination, especially for Anglophone psychology's impact on other national contexts in the late 20th century, e.g. Jevremov, Pajić and Šipka's (2007) study of the impact of Anglophone psychology on Serbian research in personality. There are examples of psychology literature analysis in other languages, e.g. Albani, Lombardo and Proietto (2014). Drawing substantive comparisons between traditions in different languages is a productive future direction for digital scholarship of the history of psychology.

[84] It would even be more appropriate to use Johanna Drucker's (2011) view of "data as capta" throughout our article. Reconceiving data as capta has some important consequences for research in digital humanities. As Drucker puts it: "Differences in the etymological roots of the terms data and capta make the distinction between constructivist and realist approaches clear. Capta is "taken" actively while data is assumed to be a "given" able to be recorded and observed. From this distinction, a world of differences arises. Humanistic inquiry acknowledges the situated, partial, and constitutive character of knowledge production, the recognition that knowledge is constructed, taken, not simply given as a natural representation of pre-existing fact" (2011; para. 3).

interesting questions and potential answers; not "complicated statistical models based on controversial assumptions" (2016; p. 215).

4.2.2 Chronologically projecting subdisciplines into term maps

Hundreds of thousands of abstracts/titles were included in our dataset. The mass of words used in them is the pool from which our terms spring, and then these terms are structured in a certain way. The structure arises from the terms' co-occurrence. How can we tell where particular groups of terms come from? In the example of *Figure 4.2* and *Figure 4.3* shown in the previous section, how can we tell where the terms in the eastern versus the western half of the map come from? For example, do they talk about mental testing or animal psychology? Where does one begin and the other end, and how do they change through time? The clusters we see in *Figure 4.2* are interesting and meaningful, but very difficult to interpret in maps based on hundreds of journals.

If the structure has any meaning, clusters of terms will tend to appear more often in the abstracts and titles of particular groups of journals. In order to scale it to the level where we can meaningfully analyze it, we need to make a selection of relevant journals that represent certain subdisciplines in psychology. To this end, we used Daniel Burgard's (2001) selection of journals of the century in psychology. Burgard, as a Psychology Subject specialist librarian, compiled a list of the "best psychology journals of the century" (p. 42). We are not sure if these journals are the best, but since Burgard used multiple criteria to make a selection of a very small number of journals, we found his selection very useful.

If we made a 'heat indicator' in our term maps based on small groups of what are perceived as excellent journals in particular subdisciplines, we would get some much needed information about the structure of the whole map. This is what we did. In the following selection of visualizations, the redder the color of a term is, the more frequently it appears in the abstracts and titles of articles in that group of journals. Doing this, we visually identify hotspots in the whole term map - and by doing it for a number of groups of journals, we glimpse an overall structure. Other than the journal projections, the figures are organized in the same way as the first map and its density visualization in *Figure 4.2* and *Figure 4.3*. We note that in the density visualization, the 'redness' of an area in a map, indicates the presence of a large number of terms in that area. In the journal projections the 'redness' of a term indicates that the term occurs relatively often in a group of journals.

Table 4.2. Journal groups according to Burgard (2001) and O'Brien (2001).

Group name	**Journal**
General	American Psychologist
Applied	Journal of Applied Psychology
	Journal of Counseling Psychology
	Personnel Psychology
Biological/Physiological	Journal of Comparative Psychology
	Comparative Neuroscience
Clinical/abnormal	Journal of Abnormal Psychology
	Journal of Consulting and Clinical Psychology
Developmental	Child Development
	Developmental Psychology
Experimental	Cognition
	Journal of Experimental Psychology*
	Perception & Psychophysics
Personality/social	Journal of Personality
	Journal of Experimental Social Psychology
	Journal of Personality and Social Psychology
Educational	Journal of Educational Psychology
	Educational and Psychological Measurement

*In the time period from 1950-1999 the *Journal of Experimental Psychology* fractured into multiple subfield specific journals. All are included in our dataset.

To project the journals through overlays, we grouped them following Burgard's scheme, although we selected just a few of the journals listed by him for ease of coding and analysis. To Burgard's categories, we also added educational psychology as a category drawing on a complementary publication, *Journals of the Century in Education* (which covers educational psychology) by Nancy Patricia O'Brien (2001). The groups and the journals that represent them can be found in *Table 4.2*.

Some journals changed names in the period: Journal of Abnormal Psychology used to be called Journal of Abnormal and Social Psychology; Journal of Consulting and Clinical Psychology used to be called Journal of Consulting Psychology. Journal of Comparative Psychology and Comparative Neuroscience used to be a single journal called Journal of Comparative and Physiological Psychology. The journals were included under the different names in their respective time periods.

We projected[85] each of the above groupings of journals into separate decade maps. These maps will be analyzed from *Figure 4.4* to *Figure 4.11*.

[85] The scale in the bottom right of every figure is not fixed in the decade maps represented in the schematic per decade visualizations on top of every figure. This was done to ease interpretability, but the reader should be cautioned against interpreting the 'presence' of each journal as being more or less constant in every decade. The later decades include many more journals and the term frequencies for each term are much larger, so the relative proportion of each journal containing the terms in their abstracts/titles is smaller. In other words, single journals have lower relative frequencies for particular terms in each of the decades following the 1950s. In the actual maps accessible through the links in footnote 79, the range can be manipulated by double clicking on the scale, and the different overlays thus generated can be inspected by the reader.

Experimental: *Cognition, Journal of Experimental Psychology*, and *Perception and Psychophysics*

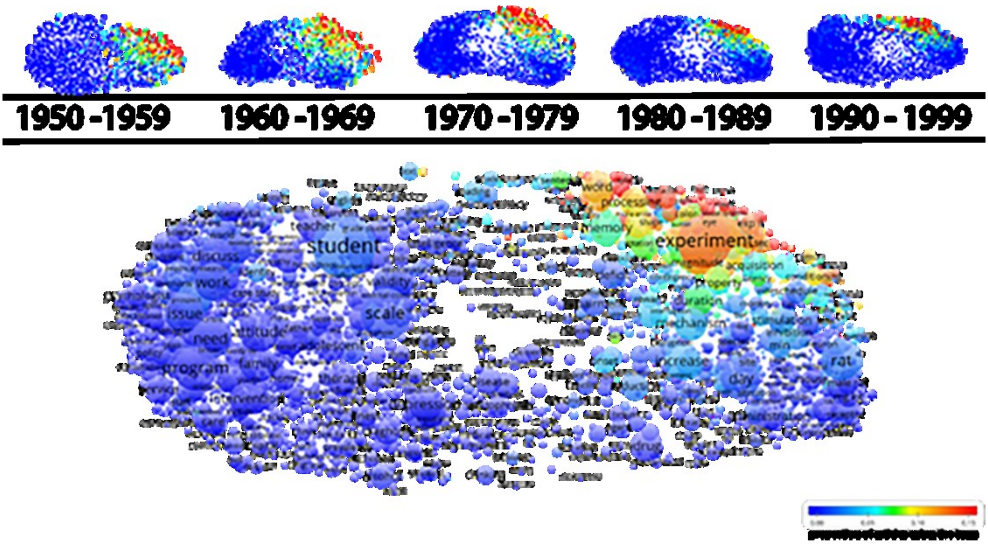


Figure 4.4. VOSviewer projections of experimental psychology

The three abovementioned journals project into the maps to generate *Figure 4.4*. As before, the big map represents the terms for the whole period of five decades while the smaller schematic images are journal projections per decade.

Looking at *Figure 4.4*, we can see that *Cognition*, *Journal of Experimental Psychology*, and *Perception and Psychophysics* grouped in the northeast quadrant of the map. Not surprisingly, they seem to have gathered around the term experiment. The further we go west and north in their grouping, the more cognitive the vocabulary becomes (with terms like "processing", "memory", and "syllable"). What we see in the top tableau is the shrinkage of coverage of these three journals. In the 1950s, they projected into the terminological territory of the whole northern part of the eastern supercluster. As we go further through the second half of the twentieth century, these three journals become less representative of the eastern half of the map. The shrinkage of the projections of particular journals through time happens with almost all of the projections we analyze. This would be even more evident if we fixed the same scale for every decade – so much so that sometimes it would make interpretation difficult. This is probably due to doubling of literature in each decade. The expansion of the literature goes hand in hand with the specialization of particular journals for smaller subfields or developing research fronts.

Biological/Physiological: *Journal of Comparative and Physiological Psychology/Journal of Comparative Psychology* and *Comparative Neuroscience*

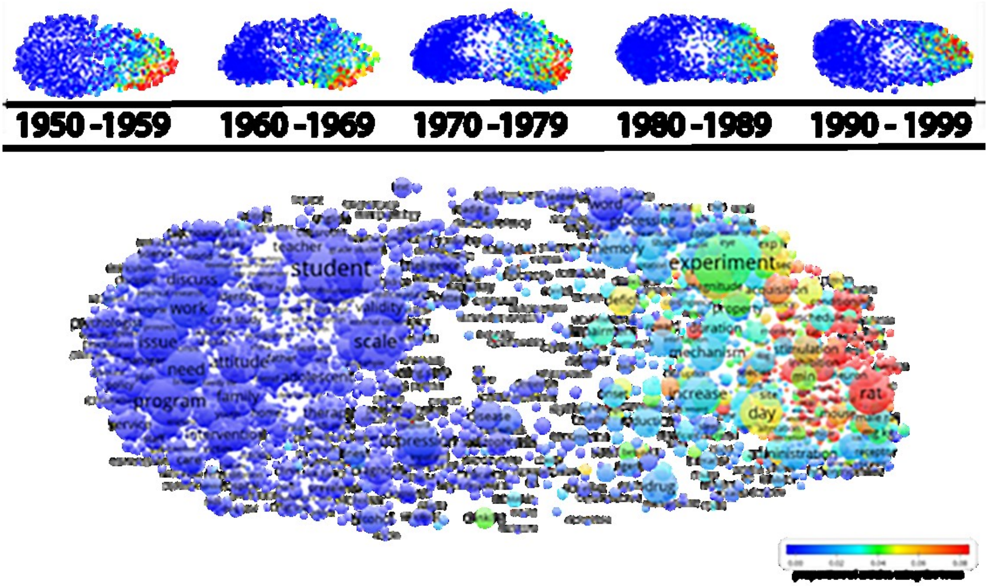


Figure 4.5. VOSviewer projections of biological/physiological psychology

Saying that the journal projection "moves" or becomes more "central" is a geographic metaphor, which means that the projection position has changed in relation to the other projections in the map, or to its position in the previous decades. Centrality is not a measure of importance for the field of psychology intellectually, or as a value judgment; it is just a comment on the relative position in the term map.

The Journal of Comparative and Physiological Psychology is projected in the *Figure 4.5*. It is situated in the eastern supercluster, and defines the east-central quadrant of the maps. In the 1950s, it projects all over the quadrant, and exhibits a similar trend of shrinkage in the coming decades. It moves from the southern position occupied from 1950s to 1970s, taking a more central position in the eastern supercluster in the 1990s. This move is due to the expansion of the terminology related to drugs in specialized journals on psychopharmacology, which start occupying the south-west part of the eastern supercluster. Furthermore, the easternmost edge moving slightly to the north in 1990s signals the rise of neuroscience - with a cluster of words like "neuron", "electrical stimulation", and "prefrontal cortex" – this is the part of projection dominated by *Comparative Neuroscience*.

Applied: *Journal of Applied Psychology, Journal of Counseling Psychology*, and *Personnel Psychology*

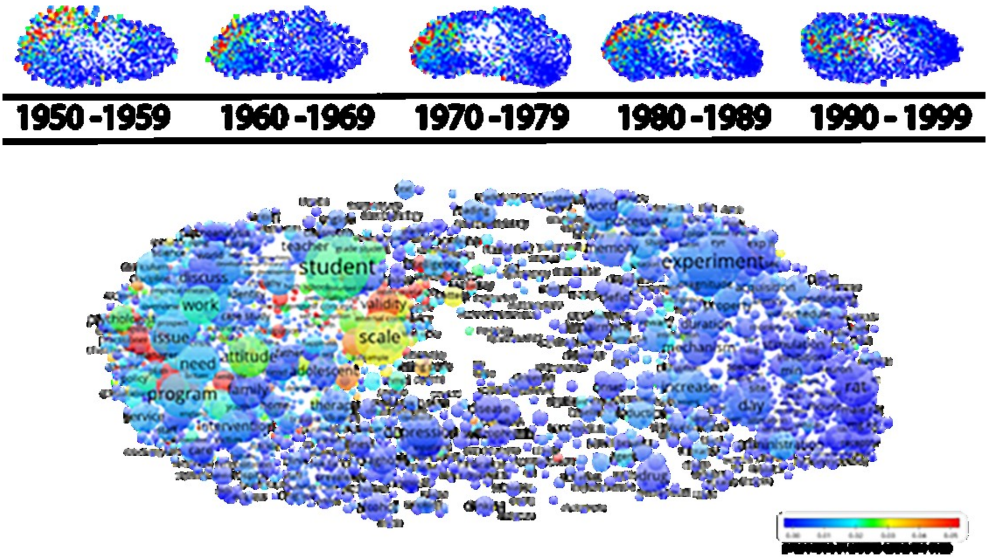


Figure 4.6. VOSviewer projections of applied psychology

These three journals project into the western supercluster in *Figure 4.6*. In 1950s, they are situated in the northwestern quadrant. In the next three decades, they shift to the southwest, occupying a central part of the western cluster. In the 1990s, they lose coherence and dissipate more. They roughly occupy the same space in the western cluster, expanding throughout its central part. In the schema for 1990s, the projection is completely absent from the northernmost part of the western supercluster - psychoanalysis shifted to occupy this space in the 1990s, which will be analyzed later.

Developmental: *Child Development* and *Developmental Psychology*

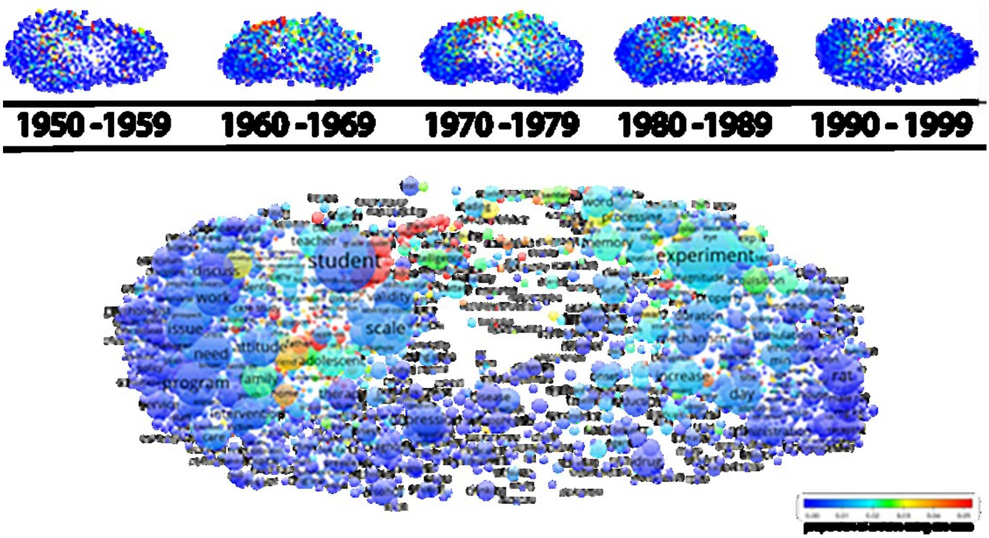


Figure 4.7. VOSviewer projections of developmental psychology

The projections for developmental psychology can be found in *Figure 4.7.* Note that *Developmental Psychology* started publishing in 1969, so the first two schemas (1950s and 1960s) are only representative for *Child Development*. In 1950s, *Child Development* projected into the western supercluster. Except for the main area in the western supercluster, there are smaller areas in the eastern supercluster that use similar vocabularies. In the following three decades - from the 1960s to the 1980s - the two developmental journals indeed straddle the northern bridge between the two superclusters. They occupy a position in the very middle of the map, jumping across the low density chasm.

We can also see how adding *Developmental Psychology* in the following decades focused the projection more and anchored it in its relatively fixed position. In the 1990s, the presence of the terms from the eastern supercluster ceases to be so prominent, and a focal axis (the red line projecting toward northeast from the term 'student' in the big overview map) can be clearly identified. The axis is actually formed by the words describing the subjects and institutions which house participants in developmental research: "elementary school child", "nursery school", "grade level", "6th grader", "kindergarten", numbered grades (from 1 to 5) and numbered graders; and the easternmost pinnacle finally bringing theoretically relevant terms: "language development" and "developmental difference." No wonder that the axis projects toward the low density chasm and the northernmost quadrant of the eastern half dominated by cognitivist experimental psychology. All in all, developmental psychology seems to be well defined in each period except the 1950s.

General: *American Psychologist*

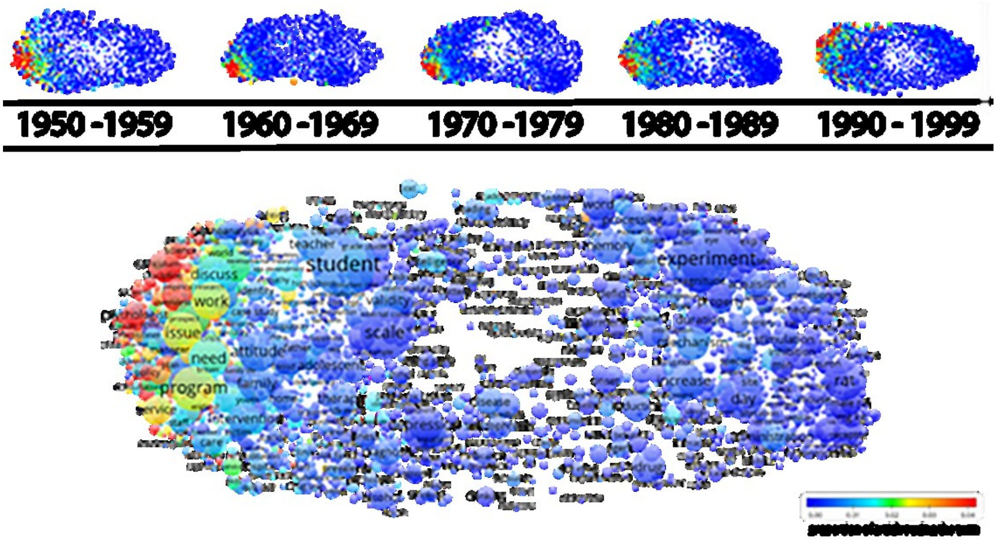


Figure 4.8. VOSviewer projections of general psychology

Figure 4.8 tells the most straightforward story. *American Psychologist* consistently projects into the same area of the map - the westernmost area of the western half of the map. This pattern does not change through the decades. Since the journal is the official organ of the American Psychological Association, this should give a clear sign that is very much in-line with the development of the APA in the 20th century and its shift from being an association of academics to an association of professionals. This shift is already in full swing by the 1950s, and its effect is consistent with the structure of our maps.

Personality/social: *Journal of Experimental Social Psychology, Journal of Personality and Social Psychology*, and *Journal of Personality*

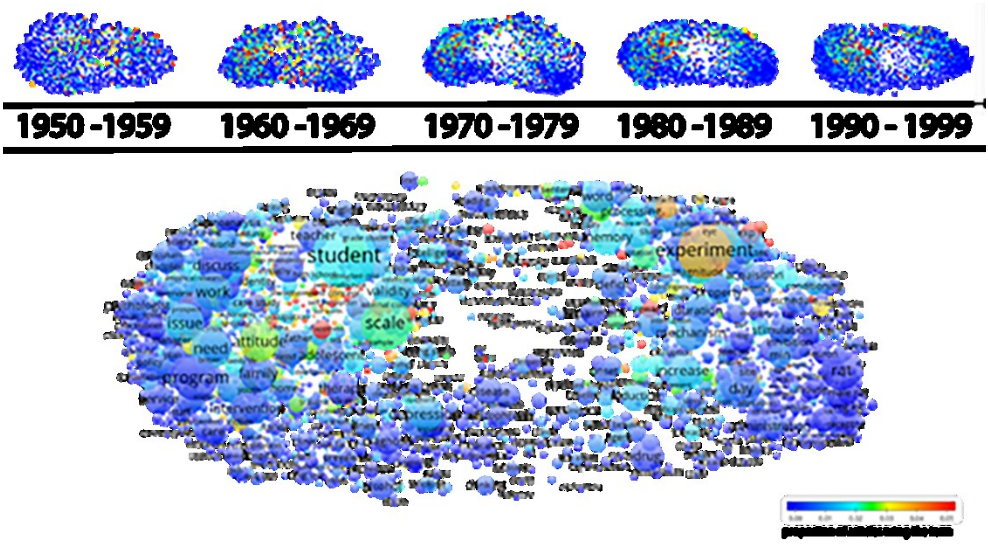


Figure 4.9. VOSviewer projections of personality/social psychology

Social and personality psychology were projected with three journals in *Figure 4.9*. Note that both the *Journal of Experimental Social Psychology* and *Journal of Personality and Social Psychology* started publishing in the 1960s, so they have no projections in the schematic map 1950-1959. Considering this, it is difficult to say how that schema compares to the other four decades, and the large overview map. Otherwise, personality and social psychology project in such a way that straddles the divide between the eastern and western supercluster, with the main concentration of terms located in the western supercluster. This concentration gains more definition with each coming decade, locating the center of the projection in the central part of the western supercluster, gravitating toward north and east. Most of the presence in the eastern supercluster is in the north-west part of the supercluster, reinforcing the idea that social and personality psychology straddle the low density gap in the middle. The map in the 1990s supports this interpretation, with scant presence of terms in the eastern supercluster and a hotbed in the west.

Educational: *Educational Psychology* and *Educational and Psychological Measurement*

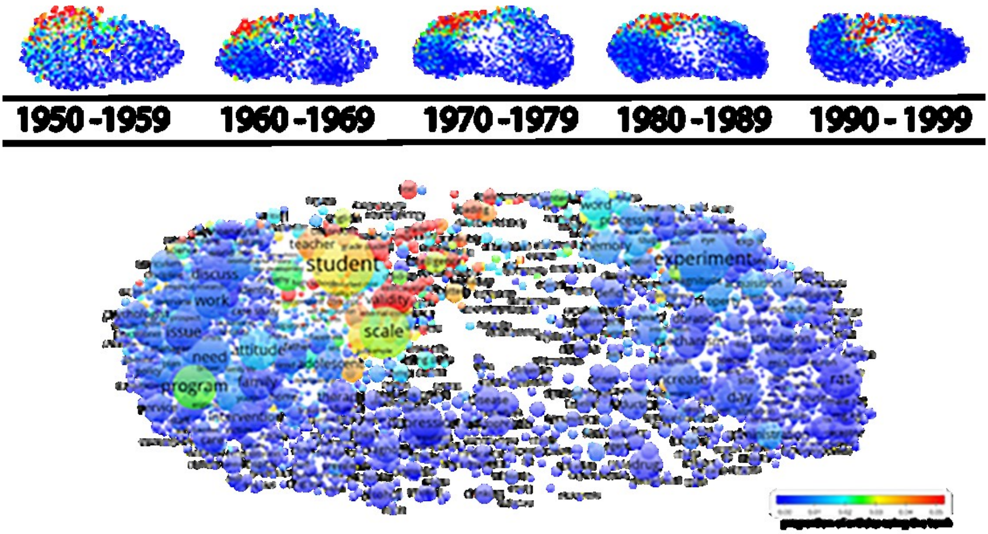


Figure 4.10. VOSviewer projections of educational psychology

In *Figure 4.10*, we analyze the projections of educational psychology. Educational psychology roughly projects in the same space developmental psychology occupied in *Figure 4.7*. This is expected and commonsensical for the whole period, providing evidence for the robustness of the map structure. Terms related to educational psychology form a cohesive unity in the northern part of the western supercluster. From the 1970s onwards, they also start to move slightly toward the eastern supercluster, but the biggest concentration of the subfield remains in the northern part of the western supercluster. The dominance of educational psychology over the whole northern quadrant of the western supercluster in the 1950s is telling of the centrality and importance of psychologists' research in the educational setting. If we link the western supercluster with Cronbach's 'correlational psychology', educational psychology, at least terminologically, encompasses a large part of psychologists' research interests among the correlationalists. There is also something to be said about the term that appears in most abstract/titles in the whole half century: 'student'. The term is used in 10,8% or 73,110 articles across the decades.[86]

[86] This is a conservative estimate, because VOSviewer separately identifies frequent collocations which include the term 'student' (e.g. "grade student," "secondary school student," but also "student's perception," or "student performance."). With a more inclusive but less accurate query in our dataset, searching for "*student*", we arrive at a bigger estimate of 82,286 entries using the term *student*.

Clinical/abnormal*: Journal of Abnormal and Social Psychology/Journal of Abnormal Psychology,* and *Journal of Consulting Psychology/Journal of Consulting and Clinical Psychology*

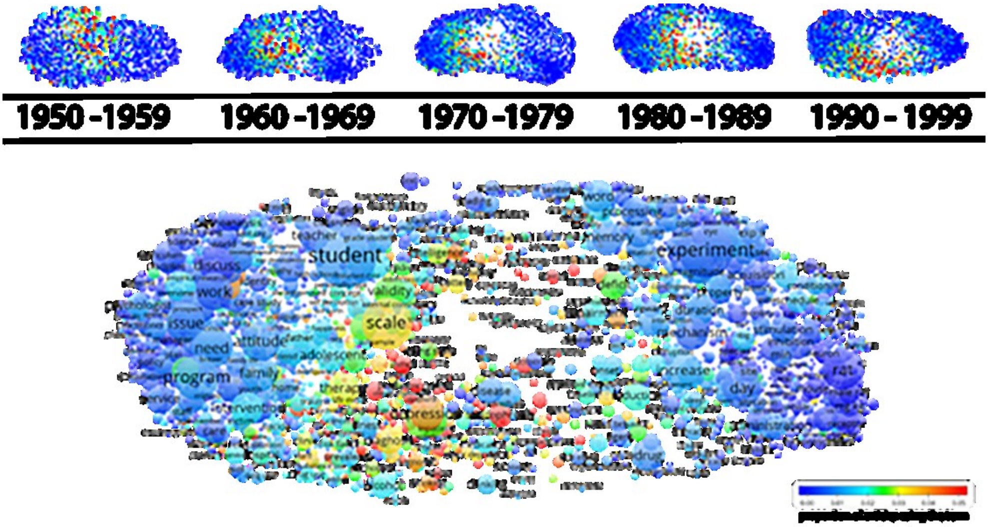


Figure 4.11. VOSviewer projections of clinical/abnormal psychology

In *Figure 4.11,* we see the projections of the two journals publishing on clinical work and psychopathology. They occupy the easternmost part of the western supercluster (the one closest to the area of low density separating the west supercluster from the eastern one). In the fifty years, the focus has shifted slowly from the north to south. Note that these two journals also straddle the chasm of low density. Their position, as in the middle of the two superclusters, becomes more predominant later. This is also related to the pharmacological perspective which slowly displaces the projection of *Journal of Comparative Psychology* - from its dominant position in the southeastern quadrant from the 1950s through the 1980s, and then its shrinkage in the 1990s. Seen in this way, the projection's shift from north to south is a shift from the terminology of psychometrics and tests toward psychopharmacology.

4.2.3 The peculiar case of psychoanalysis and psychotherapy

Even though we do not have much space to devote to psychoanalysis, its erratic behavior in the maps needs to be mentioned. In the overview map covering the whole period psychoanalysis occupies the north-western top of the western supercluster (around the prominent term 'psychoanalysis'). Around psychoanalysis, the terms form an indicative collection of psychoanalytic vocabulary: "Jung", "psyche", "countertransference", "Freud", "psychoanalytic process", "unconscious", "dream", "superego", "object relation", "narcissism", "defense mechanism", etc. Even with this short list, it is evident the terminological space is highly specified and identifies various branches of psychoanalytic schools of thought. What is peculiar is that the position of this collection of terms shifts in the period under investigation. Its shift isn't gradual like in the other cases, but drastic and it changes the spatial structure of the map.

In the overview map the psychoanalytic terms occupy the previously mentioned northern spot in the western supercluster. But in all the decade maps from 1950 to 1989, psychoanalysis can consistently be found in the south-eastern edge of the western supercluster.[87] In other words, for forty years this collection of psychoanalytic terms project in close proximity to the terms that are often used in the journals related to clinical and abnormal psychology. Then psychoanalysis separates from its consistent position and shifts to a place where it is not evident how it relates to the larger structure of the map.[88] This might be a terminological shift that recognizes the general decline of psychoanalysis (Hale, 2000) during the second part of the twentieth century. The demise of psychoanalytic language and categories as a framework for psychological research was probably catalyzed by the publication of DSM-III in 1980, and the subsequent move toward the symptom-based categorical disease model of mental illness and rise of psychopharmacology (Mayes and Horwitz, 2005).[89] This interpretation seems plausible, especially since terms like "cognitive-behavioral therapy" and "cognitive behavioral treatment" do not follow the psychotherapy and psychoanalysis exodus from the southern edge, while terminology used in psychopharmacology expands drastically. This shift in the position of psychoanalytic terms is interesting in itself, and would require an investigation of its own.

[87] If psychoanalysis is in the southern part of the western cluster for four out of five decades in our analysis, why does the overview map for the whole five decades represent it in its location from the 1990s? This has to do with literature sizes - the largest amount of literature is from the 1990s because the size of the literature doubles every decade, thus the term structure of the whole period has the closest similarity to the term structure of the 1990s.

[88] You can see this odd shift in the decade schema maps for the projection of educational psychology in *Figure 10*. There, in the 1990s, educational psychology does not occupy that much of the northern edge of the western supercluster as in the previous decades. This contraction of the projection of educational psychology from the north-west part of the western supercluster roughly corresponds to the place taken over by psychoanalysis and psychotherapy. A similar thing can be seen in the northern projection of the applied journals in the 1990s in *Figure 6* and in the slight shift of *American Psychologist's* projection from south to north in the 1990s map in *Figure 8*.

[89] We would like to thank the anonymous reviewer for pointing us toward the changes in the classifications of mental illness that were centered on the publication of DSM-III in the 1980s.

4.3 The two disciplines of scientific psychology: Lee Cronbach in 1957 and 1975

We opened the chapter's introduction with the first sentence of Lee Cronbach's 1957 presidential address, in which he likens psychology to a diverse circus. Cronbach's opinion was not only a vivid metaphor aimed at provoking knowledgeable chuckles among his colleagues in the audience. His was an incisive diagnosis of the state of the art in the discipline – a State of the (dis)Union given by its luminary. The crux of his argument was that psychology in the 1950s had been a divided discipline. There was a schism between "the Tight Little Island of the experimental discipline" and the "Holy Roman Empire" of correlational psychology "whose citizens identify mainly with their own principalities." He saw these two disciplines as "streams of thought" that have a few markers: "philosophical underpinnings, methods of inquiry, topical interests, and loci of application." He viewed the two traditions as disciplines of scientific thought, and that the "job of science" is asking "questions of Nature" (p. 671).

Our visualizations start in the decade when Cronbach identified the schism, and go on through the period when he made his second re-evaluation in 1975. Given the timeline overlap between our maps and Cronbach's diagnosis of disciplinary disunity his description provides one possible interpretation for our maps. The Holy Roman Empire is the western half - where clinical, abnormal, personnel, consulting, developmental, educational psychologies project. Then, we have the low density divide in the middle signifying the schism, and the eastern half where physiological and experimental psychologies find their location in the self-organizing pattern of co-occurring terms. Even his language fits with our geographical terminology, when he says: "The personality, social, and child psychologists went one way; the perception and learning psychologists went the other; and the country between turned into desert" (p. 673). The metaphor suddenly becomes visual, the desert being the low density area in the middle of our maps.

In 1957, Cronbach used the methodological language of psychologists, mapping the polar opposites of the two disciplines of psychology to historical predecessors: Wundt's "'experimental psychology' versus 'ethnic psychology'", Stern and Binet's "'general psychology' versus "individual psychology'", the turn of the century "'experimental' versus 'genetic' psychology", and his contemporaries' "'experimental' versus 'psychometric' psychology" (p. 672). For him, the point of the difference between the two streams of scientific psychology is an interest in two kinds of mutually exclusive conceptions of statistical variance. Correlational psychologists look at the variance arising out of individual differences; experimental psychologists at the variance arising out of two different treatments. By treatment, he means the particular kind of manipulation by the experimenter. The kind of variance that is the object of research for one group is the source of error for the other.[90]

In Cronbach's terminology, the goal of scientific research in psychology is developing better ways to control, isolate, inspect, and manipulate variance. This goal drives psychologists to develop more sophisticated methods, which often directly translates into more sophisticated statistics. In other words, better methods rigorously applied, extend psychology to different topics it can tackle and circumscribe.

4.3.1 Sophisticated methodology ≠ developed theory

Methodological sophistication translates into more topics of research, but more topics of research do not translate to theoretical sophistication, at least according to Cronbach. When evaluating how psychology was fitted to these methodological boundary conditions in 1975, Cronbach (p.116) voices his disillusionment: "Model building and hypothesis testing became the ruling ideal, and research problems were increasingly chosen to fit that mode. Taking stock today, I think many of us judge theoretical progress to have been disappointing." Moreover, he acknowledged active discontent within the ranks: "Many are uneasy with the intellectual style of psychological research […]." The research agenda of the two streams of scientific psychology have not been as rewarding as Cronbach, and many other psychologists, had hoped.

4.3.2 Cronbach's declining optimism

After his hopeful invocation of the merging of perspectives on how to conduct research in 1957, Cronbach was less optimistic two decades later. He described his view of psychological research again in the words of the reigning method - what he called the Aptitude x Treatment Interactions (ATIs). ATIs are a way of statistically analyzing the role of both treatment and individual differences variance, and how their interaction potentially produced effects different from those produced from the single source of variance. The problem, which he called his shortsightedness in 1957 (p. 119), was not only that a general statement based on treatment was misleading without taking into account relevant individual aptitudes (and vice-versa), but that the same argument could be made for the interaction effects themselves: "If Aptitude x Treatment x Sex interact, for example, then the Aptitude x Treatment effects does

[90] Note how totalizing the methodological language is for Cronbach. Is he really saying that Wundt's *Völkerpsychologie* investigated individual variance? No, but the talk of variance is the terminological space in which Cronbach can develop his theses - even if that wasn't the case for Wundt or the other predecessors he lists. Relating Wundt's *Völkerpsychologie* to the Darwinist and functionalist research traditions in American and British psychology, which focus on individual variance, does not do justice to the historical record, but it does act as rhetorical leverage for Cronbach's argument about correlational psychology being one big tradition within psychology.

not tell the story." His perspective became quite bleak when he admitted: "When we attend to interactions, we enter a hall of mirrors that extends to infinity" (p. 119).

Cronbach offered a humble conclusion for the science of psychology in the 1970s (p. 126):

> Social scientists are rightly proud of the discipline we draw from the natural-science side of our ancestry. Scientific discipline is what we uniquely add to the time-honored ways of studying man. Too narrow an identification with science, however, has fixed our eyes on an inappropriate goal. The goal of our work, I have argued here, is not to amass generalizations atop which a theoretical tower can someday be erected (cf. Scriven, 1959b, p.471). The special task of the social scientist in each generation is to pin down the contemporary facts. Beyond that, he shares with the humanistic scholar and the artist in the effort to gain insight into contemporary relationships, and to realign the culture's view of man with present realities. To know man as he is is no mean aspiration.

4.3.3 Who are the correlational psychologists?

We should also note that Cronbach's naming convention for the schism – experimental and correlational – shows a certain bias. As one of the people pushing for a correlational scientific psychology in educational but also other "applied" settings, Cronbach found much to gain by gathering all the disparate non-experimental psychologies under the banner of the correlationists. Keeping with his metaphor, the Holy Roman Empire might have been disunified, but at least it had an emperor. The possibility that the professionals and scientists working in the disparate educational, clinical, organizational, counseling, etc. psychologies might have disagreed with such a designation became invisible – they were all gathered around a single (statistical) conception of their research object. We don't have to think hard to provide historical examples that correlational psychology wasn't a wholly uncontroversial description of non-experimental psychology for some psychologists in the second part of the twentieth century. One of the well-researched examples is humanistic psychology as the "third way" and its different conception of what a science of psychology should be. A good example of such a different conception of science is Abraham Maslow's, which he described in his *Psychology of Science* and other publications (Maslow, 1966; Kožnjak, 2016).

Whether we call it correlational or some other name, the integration of the east and west, according to Cronbach, had failed. Early 21st century evaluations agreed with him (Borsboom et al. 2009). But what role has this failure played in how psychology developed into a discipline in the second part of the 20th century? Our thesis is that the failure to integrate structured scientific psychology around purely methodological lines. The mass of theories and models – the substantive content of the discipline – just followed suit and kept expanding alongside the lines laid out by the sanctioned ways of doing research. In a nutshell, the theories a psychologist wanted to test were never in the front seat – they came and went as they conformed to the ways she could research them. The names of what psychologists call “constructs” could multiply indefinitely, allowing for more and more research that never fed back into larger developed theories, as long as the constructs fulfilled the methodological criteria put before them.

4.4 Psychology’s methodological infrastructure

Borsboom and colleagues (2009) gave an expanded analysis of the problem of approaching variance in different ways. Their conclusion was even less optimistic than Cronbach’s in 1975. They concluded that the integration of these two strands of research in psychology “is a dreamed route of progress that is really a dead end street” (p. 94). We are not primarily interested in how to solve this lack of integration, but to explain how it was sustained for so long despite vocal criticism from prominent psychologists and historians. Borsboom and colleagues’ have an idea of what kept it vital (p. 82):

> Theoretical camps professionalize in such a way that a given method is sanctioned, findings that employ the method are publishable when review by the professional in-group, and the body of published findings sustains the theoretical approach, including the careers of those who espouse it.

In other words, methodologies are the collective choices of communities and those choices act in sustaining differences between various approaches. Choice of methodology is a more charitable name for what Kurt Danziger (1985) calls methodological imperative, or in an even more negatively charged tone, methodolatry (e.g. Bakan, 1966). In Danziger’s (1990; 1997) and Winston’s (2005) view, the role of the meta-language and the methodological discourse is to frame the institutional and substantive boundaries of psychology. By meta-language we mean the basic vocabulary of psychologists in the second part of the 20th century, in which disparate communities of psychologists have been educated in graduate schools, and which appears in textbooks and APA manuals. It includes the vocabulary of variables (independent, dependent, mediating, intervening), operational definitions (constructs, measures), and research designs (randomized experimental trials, longitudinal and transversal correlational studies, quasi-experiments). According to Danziger (1985, p. 10), this “widely accepted methodology”[91] in fact involves theoretical commitments and “this shared commitment … provides the basis for effective intra-disciplinary communication.” The supposed theoretical disunity is manifest, but underlying it is a methodological common language suffused with the minimal

[91] The methodology, alongside the signposts of the meta-language we have included above, also includes inferential statistics and internalized philosophies of science of various psychologists (namely, operationism, see Green, 1992b; Feest, 2005).

theoretical commitments of the psychological sciences as a whole. To put it straightforwardly, doing "correlational," "experimental" or some hybrid of the two is already circumscribed by a tradition of thinking that presupposes that the object of research can be found in aggregate statistics and the variables describing them. Cronbach's gathering of the patchwork of non-experimental psychologies under correlational psychology becomes an example of how totalizing the methodological language can be. The psychometricians, in their conclusion, state that integration is impossible in principle (Borsboom et al, 2009). We would add that the correlational/experimental distinction is not even an issue being addressed in research practice, considering that psychology expands (progresses) as usual *because* this methodological consensus exists.

Hank Stam puts it bluntly (2004, p. 1262): "Calls for unification, no matter how well articulated, will likely fall on deaf ears since there are already deeply entrenched positions in the discipline that are supported by the implicit unity of method and framework." Coming back to our analysis of the term maps - we show that psychological literature exhibits a stable unchanging structure in five decades - despite the internal squabbles in many of its subdisciplines. Chris Green (2015, p. 210) briefly sketches these changes in psychology: "the integration of the American Association for Applied Psychology with the APA in the 1940s; the simultaneous emergence of the Boulder Model; the rise and fall of "third wave" humanism; the partial retreat of behaviorism as a theoretical basis for psychology and its reformulation as a leading basis for various therapies; the splintering off from the APA, first, of the Psychonomic Society and, later, of the American Psychological Society; the appearance of the Psy. D.; the rises of cognitivism, computationalism, evolutionism, neuroscience, and so forth." All these supposedly tectonic changes (or at least earthquakes) have happened, alongside the massive expansion of the number of researchers, practitioners, and subdisciplines; yet the field exhibits a stable structure if analyzed as a whole.

We argue that the minimal methodological commitment with its theoretical baggage (whether we call it methodological imperative, meta-language of variables, or just methodology) is the robust structure of the superclusters. The psychological sciences, then, are a "collection of generalizations that describe relations among classes of variables, the kinds of relations and the kinds of variables being predetermined by the methodology" (Danziger, 1985, p. 11). Growth and expansion in the period from the 1950s is just in magnitude (of data being analyzed and generalizations being made); the structure of the knowledge stays the same. Or in the pessimistic tone of both Cronbach's 1957 article and Danziger's writing on the state of research in the second part of the twentieth century: Generalizations have been amassed, data gathering and manipulation has grown in sophistication, but theoretical progress is almost nonexistent.[92] As Stam puts it (2004; p.1261): "[P]sychology proceeds through the multiplication of entities without ever committing itself to the reality (or lack thereof) of the objects it so constitutes."

Another way of asking why the structure of the literature is so stable is to stop equating the literature with the discipline, as we have done up to now. If we follow Green's (2015, p. 210) incisive observation that "[p]sychology has been a hodge-podge from its very creation, and none of the various efforts to create a theoretically unitary discipline out of that miscellany has ever been remotely successful," psychology throughout the 20th century has been a discipline assembled out of incompatible parts. These incompatible parts, however, at least from the 1950s onwards, produced a literature with a stable structure. Then the question is: why haven't these incompatible parts of the discipline balkanized its literature?

The answer, we argue based on the stable and interpretable patterns in our maps, is because the respective methodologies *in* each of the incompatible parts *did* remain stable throughout the second half of the 20th century. On one end, the east of our maps, the conglomerated methodology provides an "objectification of subjectivity", a set of mechanical rules for scientific thinking (Gigerenzer et al, 1990., p. 84). The widespread use of descriptive quantification accompanied by appropriate inferential statistics, and the research designs this usage presupposes, provide the scientific backbone "experimentalists" feel they can rely on. On the other hand, the western side fills its lack of systematics and theoretical justification by placing its bets on the scientific rigor of its shared methodology. Psychoanalysis - and other large theoretical and metaphysical systems - can be left in the dustbin of history because the scientific rigor of psychology's methodology has taken their place; and that scientific rigor is not only applicable in the clinics, schools, and hospitals where psychologists ply their trade. The rigor shares a family resemblance to the laboratories that produce psychological knowledge. And we arrive at the true promise of the Boulder model: The scientist-practitioner has a discipline to call home.

Our argument that methodological uniformity kept the centrifugal forces at bay runs the risk of being read as a simplification, the kind of simplification to which intellectual historians are prone when trying to describe historical change. Methodology becomes an abstraction – a hidden cause – that kept experimentalists and correlationists (and all those others made invisible by Cronbach's two designations) within the same discipline. We *do not endorse* such a view. Methodology and the rules of "good" research become what they are through institutionalization – in journal policies, funding structures, graduate education, writing manuals, textbooks, and handbooks. There is no decoupling of the intellectual from the social.

[92] The charge for the lack of theoretical progress in psychology is often made by falsificationists. A good example is a discussion of publication bias in which Christopher Ferguson and Moritz Heene (2012) call theories in psychology "undead" – even when they're disproven, they remain in the psychologists' canon.

All this forms a rich tapestry of historical change (or continuity), usually described by detailed microhistories. Our aim was to look at that coalesced institutional inertia from the bird's-eye view as represented by the patterns in literature. Our bibliometric analysis was not done with the intent of invalidating social and extra-intellectual explanations of historical change, but to frame them and articulate them on the macro level. As we said at the outset, we take psychology's literature as a proxy for the discipline – but the relationship between the literature and the discipline producing it is much more complicated than that, and requires further explication.

4.5 Conclusion

Our conclusion might seem contradictory. If the two disciplines of scientific psychology cannot build on each other in practice and in principle, why is their collective structure in the period of fifty years stable? Let's shortly go through the argument in the article to see where the contradiction comes from, and how to resolve it.

We have collected a large set of abstracts and titles of published literature in psychology. Then, we have extracted a set of terms out of those abstracts/titles, calculated their co-occurrences, and visualized the co-occurrence relationships between terms, creating the term maps. We have made comparable maps for every decade in the period under investigation. After that, we have projected sets of journals that are the flagships of their subdisciplines into the term maps, and described the structure of the terms in light of how these different journals occupy space. This procedure identified a stable structure of the literature in the psychology from 1950 to 1999. The only large structural deviation that remains conspicuous is psychoanalysis. The structure included two large superclusters, the eastern and the western one. The eastern one included experimental and physiological psychology. The western included various subclusters of educational, social/personality, clinical, and what are traditionally called applied psychologies.

We have connected this structure to a description of psychology made by Cronbach in 1957, of the science being constituted by two historically distinct streams of method. Cronbach, in the 1950s, hoped for a unification of these two streams. That such a thing never happened is confirmed by Cronbach's reevaluation in 1975, a psychometric analysis of the state of the art in 2009 (Borsboom et al, 2009), and the structure of our maps.

How do we answer the contradiction between the stability of the term maps' structure through time and the inability for the two separated superclusters to integrate? Our thesis is that the structure represents the methodological metalanguage of psychology.[93] In Danziger's words, the degree of coherence is caused by the methodologically embedded principles that define the limits of acceptable research and theorizing (1985, p. 10). These principles are the methodological standardization that led to viewing the object of research of psychologists as the manipulation, analysis, and prediction of variables. The method not only defines the universe of potential answers to the questions psychologists are interested in, but more importantly, it defines the boundaries of what questions one can ask to begin with.

In this way, we provide bibliometric evidence for the theses of Danziger, Stam, and Winston. Methodological standards have facilitated the massive expansion, both in size and number, of journals in which psychologists publish in. This has allowed for growth of the discipline. That growth, however, has been checked in the last few years by the looming replication crisis (see Open Science Collaboration, 2015). Maybe instead of a fragmentation into different research superclusters which is suggested as one possible scenario by Green (2015), or unification, as announced by so many others, we will actually see a great culling of viable research areas in psychology. With that thought, our historical bibliometric analysis is a call for recognizing that psychology's theoretical and metaphysical content has been reduced to a barren methodology. In an extreme conclusion that could be drawn from our analysis, there are no theoretical systems to discuss in scientific psychology in the 21st century. There is just methodology.

Our conclusion is somewhat familiar even if we garb it in digital humanities and data-mining of literature. The terms in our maps become a proxy for the content of psychology - with its few theories, many models, descriptions of participants of research as students or various animals or children, names of particular methods and research designs, and professional discussions and negotiations of psychologists about their fields on the meta-level. The structure of those terms - the thing behind their stable spatial orientation to each other - is a methodological consensus among psychologists about the kind of research that produces psychological knowledge. The consensus allowed for an extremely efficient expansion of psychology in the second part of 20th century to everything from rats in mazes to moral behavior. All these topics remain nominally disunified under the name psychology, while they are actually serviced and sanctioned by a uniform methodology and underlying philosophy of science. The real problem is not disunity or disintegration; rather, it is the lack of a framework for discussing the way psychologists do research that goes beyond the idea that if psychology is to be scientific, it must be based on our currently accepted methodological standards. The methodological straightjacket contained the disintegration of psychology, but it is debatable whether it permits the development

[93] At this point in my research, I used a different name for what I called the institutionalized conventions of scientific psychology in the thesis' introduction. Since this chapter was published already, I did not want to retroactively rewrite the above paragraphs, so keep in mind that 'metalanguage' more or less stands for conventions of scientific psychology.

of a science that can ask new questions and provide interesting answers. The increasingly vocal jury of replicationists and crusaders for 'robust' psychological science in the 21st century is still out,[94] and we can't tell what the verdict will be for the flood of psychological knowledge gathered over the course of the 20th century – hot air or a progressing and branching empirical science of psychology? And if the latter: Per whose standards?

[94] The discussion is still fresh in the psychologists' blogosphere, and the last large chapter in the vitriolic online debate at the time of writing of this article was a surge of comments to Susan Fiske's (2016; the draft can be found in Gelman, 2016) leaked early draft of a column accusing some psychologists of "methodological terrorism" in the APS Observer. For a polemical answer, see Andrew Gelman's (2016) blog post. For an insightful journalistic piece covering the debate, see Jesse Singal's text (2016).

Chapter 5. Psychologists psychologizing scientific psychology

> This book is borne out of what I can only describe as a deep personal frustration with the working culture of psychological science. I have always thought of our professional culture as a castle – a sanctuary of endeavor built long ago by our forebears. Like any home it needs constant care and attention, but instead of repairing it as we go we have allowed it to fall into a state of disrepair. The windows are dirty and opaque. The roof is leaking and won't keep out the rain for much longer. Monsters live in the dungeon. (Chambers, 2017. p. IX)

Chris Chambers paints a dark picture in the opening paragraph of his book on the most recent crisis in psychology.[95] Under the title *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*, Chambers offers a *tour de force* of issues that have cropped up and disturbed researchers in psychology in the past years. Organized around the tongue in cheek metaphor of psychologists' seven deadly sins, Chambers' manifesto is an informed criticism of what went awry with psychology as a science: The institutionalization of biases, dubious flexibility in the usage of statistical procedures, lack of transparency, outright fraud, and systemic perverse incentives. He concludes the book with a chapter calling for reform, arguing for the institution of a new publishing practice called registered reports and, with it, improved best practices for research in psychology.

I picked out Chambers' account of the crisis in psychology, among other prominent voices in the reform movement,[96] for a specific reason. It offers a deeply personal view on the current crisis, but that personal view is used as a springboard for normative, methodological, and statistical discussions. Chambers' view succinctly exemplifies the following: For many psychologists, be they reformers or fellow travelers, what is currently at stake are the norms of good science, the viability of their research programs, and in the long run, their survival in the competitive funding structures of science at large (see Green, 2018). The reformers among psychologists argue for robust enforcement of what they perceive as the norms of science. In doing so, they provide

---

[95] For a through discussion of a number of crisis episodes in psychology, see the special issue of *Studies in History and Philosophy of Science* (Sturm & Mülberger, 2012).

[96] I follow Maarten Derksen (submitted) in calling it a 'reform movement', and the psychologists, methodologists, and statisticians that are criticizing research practices the 'reformers'. They are a wide group of researchers calling for more replication studies and "increased rigor" of psychological research. For an overview, see Shrout and Rodgers (2018).

an explication of what is ‘healthy’ research in a very particular way: By employing a reconstruction of what science is, coupled with a psychologically-informed view of who scientists are. From these two usually implicit views of *what* is science and *who* scientists are (or in some accounts, should be) as people, they criticize current research practices and provide solutions. In this chapter, I will make these usually implicit views on science and scientists explicit.

In order to investigate the reformers’ conception of science and scientists I will use an analytical category called “indigenous epistemology” developed by the historian of psychology Laurence Smith (1986). I will attempt to reconstruct the indigenous epistemology of the reform movement and critically contrast it to the kind of philosophy of science psychologists use when discussing science reform. My aim is to constructively criticize the reform movement by making explicit how psychologists psychologize scientific psychology, and by extension, to point out where such psychologizing needs more conceptual work, especially when it uses the work of philosophers of science. The reformers in their writing tentatively subscribe to various positions on ways of knowing and functioning of the science system. There is a big discrepancy in how the reformers discuss those epistemological questions, and how such questions are debated among historians and philosophers of science. So much so that the reform debates seem to be completely out of tune with contemporary history and philosophy of science. In this chapter, I try to map these disparate discussions about similar topics, and in doing so, signpost how to move the discussion forward.

In section 5.1, I will describe what indigenous epistemologies are in practice using two historical examples: The first example is neobehaviorism for which Smith proposed the concept of indigenous epistemology in the first place, and the second example is interpreting Abraham Maslow’s humanistic psychology as a kind of indigenous epistemology. In section 5.2, I will describe my reconstruction of the indigenous epistemology that is dominant in the current reform movement. I call the one the reformers use the indigenous epistemology of irrationality. In section 5.3, the reformers’ indigenous epistemology of irrationality will be criticized. In section 5.4, I will discuss some implications of psychologists’ indigenous epistemologies being a kind of naturalized epistemological position. Naturalized epistemologies are a controversial set of positions among philosophers in the 20th century, but this is almost never acknowledged in the reform discussions. Finally, in the chapter’s conclusion, I will try to offer some suggestions on how to move the reform debate forward and maybe change its parameters.

5.1. Indigenous epistemology as an analytical category

5.1.1 Scientists are rats in a maze searching for truth

In 1986, Laurence Smith published a lengthy historical reanalysis of the relationship between behaviorism and logical positivism in the first part of the 20th century. Smith’s work was an answer to the then standard view that there was an alliance between logical positivist philosophy of science and the behaviorists' research programs, like the research lines of the most prominent neobehaviorists Hull, Skinner, and Tolman. Smith wrote a history strongly arguing against such an interpretation. According to him, neobehaviorists used their own research to make sense of their science:

> [F]rom the beginning of their careers, the principal neobehaviorists – Tolman, Hull, Skinner – developed views of science that evolved out of and in close parallel with their presuppositions about the nature of organismic behavior. [...] All along, and in their separate ways, they were striving to develop naturalistic behavioral epistemologies that would encompass all forms of knowing, from that of the laboratory rat to the highest forms of human knowledge – including science itself. (Smith, 1986, p. 19)

These "indigenous epistemologies" were fellow travelers to logical positivism, not its applications. Interpreting them as just parts of the dominant philosophical view on science doesn't do them justice, considering they were actually worldviews built as epistemological extensions of the psychologists' empirical work. A striking example of this psychologism is the metaphor of the maze in Tolman's view of science: "Tolman held the world to be a complex, richly articulated maze that comes to be known in varying degrees by rats, ordinary humans, and scientists alike by means of exploratory activity" (Smith, 1986, p. 136).

To learn about behavior, Tolman researched rats navigating mazes and famously postulated that they develop cognitive maps to do so. Rats in mazes were a powerful metaphor for him, and he readily used it to explain the behavior of scientists: "[Cognitive maps] could serve as effective guides for action in an ambiguous and changing environment" and consequently "…science was to be understood not in logical terms but in psychological terms – or, to be more exact, in the spatial terms of his cognitive behaviorism" (Smith, 1986; p. 137). For Tolman, all human knowledge was purposive behavior. Science wasn't a system of analytical and synthetic propositions of the logical empiricists, but a psychological system of spatial relations.

Another important aspect of indigenous epistemologies of the neobehaviorists was that they were naturalized. Naturalism in epistemology (for a comprehensive overview, see Rysiew, 2017) is a contentious issue for philosophers, especially in the second part of the 20th century. It subsumes a range of possible positions, so Rysiew considers it as "more a movement or a general approach to epistemological theorizing than it is some substantive thesis (/theses)." Naturalized epistemology is any epistemology that takes "the attitude that there should be a close connection between philosophical investigation—here, of such things as knowledge, justification, rationality, etc.—and empirical ('natural') science" (Rysiew, 2017). Consequently, if a late 20th century psychologist develops an indigenous epistemology, it will be naturalized. Why that is so and what are the consequences of such psychological naturalization, I will discuss in section 5.4.

Smith also reconstructs different indigenous epistemologies for the other two neobehaviorists – Skinner and Hull. Instead of describing those here, I will introduce another historical example of a prominent psychologist developing an indigenous epistemology: Abraham Maslow describing science through his theory of personality and motivation.

5.1.2 Carving an epistemological middle way: Maslovian science

Abraham Maslow's humanistic psychology was also a type of indigenous epistemology – a description of the science system and scientists as subjective actors that arose directly (or, even, that was developed concurrently) with Maslow's thinking about motivation and personality. For our discussion, the central features of Maslow's theory of motivation[97] are relevant, "its universalism and antirelativism, its biological essentialism, and its explicit connection between the healthy person and the healthy society" (Weidman, 2016a; p. 114). Maslow approached science as a social manifestation of the internal dynamics of human nature and developed his view in the 1966 book *Psychology of Science: A Reconnaissance*. He was also a reformer, but one very unlike today's reformers, because he was arguing for loosening up what he perceived as the too strict standards inherited from behaviorism. He wished for a more inclusive humanistic psychology. Maslow's indigenous epistemology was constructed as a direct criticism of the research programs of the neobehaviorists Laurence Smith was writing about.

Maslow's goal in reforming science as a system was to redesign it in such a way that it allows for the full actualization of the "empirical attitude" even non-scientists could take, by virtue of being human beings. He defined the empirical attitude as "looking at things for yourself rather than trusting to the a priori or the authority of any kind" (Maslow, 1966, p. 135). It was the kind of healthy skepticism anybody could internalize, from "a child [...] watching an anthill" to "a housewife comparing the virtues of various soaps" (1966, p. 136).

In order to achieve a healthy and creative expression of the empirical attitude, scientists needed to integrate dichotomies that could be pathological if taken to extremes. Maslow named and represented these dichotomies in different ways that meshed the social level of the functioning of science as a system of norms and institutions and the individual functioning of scientists as human beings. He gave this dichotomy of what he called "Two Sciences" (p. xv) many names: Mechanistic and

[97] For a detailed elaboration, see Maslow (1943; 1954), and Kožnjak (2017; p. 261-264). Kožnjak also gives an extremely interesting account of Maslow's interaction with Kuhn, and discusses what would be the historical and philosophical implications of psychology of science as a self-standing discipline (comparable to sociology of science). Contrasting Kožnjak's views on Maslow's psychology of science to my view of Maslow's psychology as a naturalized indigenous epistemology is extremely interesting, but goes beyond this chapter.

humanistic science, safety science and growth science, spectator knowledge and experiential knowledge, simpleward and comprehensive science, abstractness and suchness meaning, controlling science and taoistic science, desacralized and sacralized science, means-centering and problem-centering. Using the various names he discussed the dichotomy in different ways, spelling out implications for scientists as social actors, their object/subject of research, science education, human happiness; and ontological, epistemological, and ethical consequences of the two extreme sides. Fundamentally, all those implications were drawn from Maslow's view of science as the product of human nature.

Science was a type of institutionalization of humans' cognitive activities (p. 21-22), or "a technique with which fallible men try to outwit their own human propensities to fear the truth, to avoid it, and to distort it" (p. 29). These activities were prompted by cognitive needs that are "instinctlike and therefore defining characteristics of humanness (although not only of humanness), and of specieshood" (p. 20). The crucially interesting feature of these cognitive activities for Maslow's analysis of science was that they can be instigated by fear and anxiety or, on the other hand, proceed without fear, courageously: "Behavior, including the behavior of the scientist, can be seen in simplest schema as a resultant of these two forces, that is, as a mixture of anxiety-allaying (defensive) devices and of problem-centered (coping) devices" (p. 22). For Maslow, the same mechanisms and goals can be "neuroticized" or "healthy" (p. 30), and the way individual scientists resolve them has a consequence for the kind of science they produce. In his view, psychological health of self-actualized individuals wasn't just morally good, but also a requisite for scientific creativity: "This ability to be either controlled and/or uncontrolled, tight and/or loose, sensible and/or crazy, sober and/or playful seems to be characteristic not only of psychological health but also of scientific creativeness" (p. 31). The epistemological consequence of these opposite ways of knowing was far-reaching: "The merely cautious knower, avoiding everything that could produce anxiety, is partially blind. The world that he is able to know is smaller than the world that the strong man can know" (p. 32).

Seeing Maslow's psychology of science as an indigenous epistemology provides us with three things. First, it shows that developing indigenous epistemologies was not a quirk of neobehaviorism. On the contrary, I would argue that it is a necessary consequence of the psychologists' subject matter – if one tries to construct a scientific view about the psychology of individuals, that view will by definition encompass the psychology of the scientist, especially when the psychologist is prompted to reflect about their own scientific practice by a perceived crisis. Second, since Maslow's theory of motivation is "instinctlike," thus biological, it necessarily makes his indigenous epistemology naturalized i.e. rooted in human biology and psychology. Third, it shows how for American psychologists in the second part of the twentieth century, the human mind was a concept that bundled up multiple aspects of contemporary culture, politics, and science. Maslow's account of the scientist's mind – especially the creativeness possessed by a "strong man" – sounds very similar to the model of the open mind of the human, scientist, and model American citizen that was gaining traction in the salons and research centers where the new cognitive science was coming into existence (Cohen-Cole, 2014, p. 141-214). In other words, during the 1960s, a kind of discourse developed in psychology that could sustain a scientifically descriptive account of human nature as an object of study, a normative ideal for the conduct of scientists, and a model of the mind of a good citizen in a pluralistic society.

The example of Maslow shows that using indigenous epistemologies to analyze psychologists' understanding of science is fruitful. It also shows how a particular explication of an indigenous epistemology can be wielded as criticism of the *status quo* in the psychologist's discipline – Maslow's *Psychology of Science* was not just a description of science, it was also his treatise on how to reform what he perceived as rigid behavioristic psychology in order to make it conform to his humanistic visions of science and human nature. In the rest of this chapter, I will illustrate how a different naturalized indigenous epistemology is used to call for a different kind of reform in the replication crisis debates.

5.2 Reformers' indigenous epistemology of irrationality

The fundamental issue for psychology's reformers are biased researchers. Late 20th century psychological theories of reasoning tell us that human inferences and thinking operate in a biased way. Even when humans try to be rational and objective, we do not always succeed. And this is not only true for humans "in the wild," but also for scientists. Scientists' thinking is exposed to the same kind of biases as that of other people – they are not excluded by virtue of their status or training. What's even more worrisome, scientists are in the business of knowledge production. If their practice of knowing is biased, what does that mean for the knowledge they produce? How should an irrational actor produce rational arguments and theories?

According to the reformers, where scientists should have an upper hand over other humans is in devising strategies for insulating thinking from bias, and maintaining institutions for facilitating and enforcing those strategies. The interaction between individual biased thinking and social structures of science failing to control for it is the source of criticisms advanced by the reform movement. Marcus Munafò and colleagues (2017, p. 1) put it in the following way in their manifesto for reproducible science:

> A hallmark of scientific creativity is the ability to see novel and unexpected patterns in data. John Snow's identification of links between cholera and water supply, Paul Broca's work on language lateralization and Jocelyn Bell Burnell's discovery of pulsars are examples of breakthroughs achieved by interpreting observations in a new way. However, a major challenge for scientists is to be open to new and important insights while simultaneously avoiding being misled by our tendency to see structure in randomness. The combination of apophenia (the tendency to see patterns in random data), confirmation bias (the tendency to focus on evidence that is in line with our

> expectations or favored explanation) and hindsight bias (the tendency to see an event as having been predictable only after it has occurred) can easily lead us to false conclusions. Thomas Levenson documents the example of astronomers who became convinced they had seen the fictitious planet Vulcan because their contemporary theories predicted its existence. Experimenter effects are an example of this kind of bias.

Scientific research, for these psychologists, is essentially data production with interpretation. The problem, for reformers, is the "with interpretation" part – because the biased nature of human cognition leads to analytical flexibility. Scientists should act rationally and try to insulate their thinking from their bias. This, however, presupposes two things: A certain idea of what it means to 'act rationally' and secondly, how to generalize 'acting rationally' in order for it to manifest on a large scale, on the level of the system of scientific institutions. I will discuss both in this section.

In the post-World War II period, American human sciences were a melting pot for a new brand of rationality. In *How Reason Almost Lost Its Mind*, Erickson and colleagues (2013) make a historical case for a break in the ideal-type of rationality in the 20th century from its Enlightenment ideal. The ideal-type of Cold War rationality was (p. 3) "formal, and therefore largely independent of personality or context", so "it frequently took the form of algorithms – rigid rules that determine unique solutions – which were moreover supposed to provide optimal solutions to given problems, or delineate the most efficient means toward certain given goals (taken, in this instance, for granted)." It also presupposed that "complex tasks and episodes were analyzed into simple, sequential steps; the peculiarities of context, whether historical or cultural, gave way to across-the-board generalizations; analysis took precedence over synthesis" (p. 4). What the various advocates of this new rationality hoped for was that "the rules could be applied mechanically: computers might reason better than human minds" and for the historians and philosophers writing the book this made obvious the fact that Cold War rationality as an ideal-type had historical origins: "[O]n the one hand, in the mathematics of algorithms, linear programming, and game theory; on the other, in the theory and practice of economic rationalization" (p. 4).

These developing views on rationality impacted psychology in a particular way, the authors go on to argue. "[P]sychology itself, previously a contributor to the exchanges of ideas and individuals that defined the field of Cold War rationality, began to dissolve it by redrawing the boundary between rational and irrational" (p. 22). In the period after the Cold War "psychological research purported to show stubborn, widespread biases and inconsistencies in actual human reasoning" and "[i]n the process, the science of psychology shouldered the responsibility for explaining deviations from rationality, leaving its definition to other fields" (p. 22). In other words, the authors argue that rationality, for psychologists since the 1980s, had fractured into different camps. The debate - the one whether psychologists should follow Kahneman and Tversky in the investigation of how "human beings often and systematically violate norms of rationality that derive from formal logic, probability and decision theory" (Sturm, 2012; p. 66) or the defenders of "bounded rationality" like Gerd Gigerenzer - has become known as the "rationality wars" of the 1990s (Samuels, Stich, & Bishop, 2002; cf. Sturm, 2012).[98] What was fundamentally at stake was the extent and manner of incursion of psychological descriptions of reasoning into the descriptions of reasoning produced by formal logic and probability theory, or in other words, how thoroughly are we warranted to psychologically naturalize rationality?

How did this fractured rationality enter the debates about psychology's replication crisis in the 2010s? To answer that question, I identified a group of prominent papers published by the reformers in a larger set of papers citing the review of research on the confirmation bias by Raymond Nickerson (1998). In that way, I identified a few highly relevant and impactful papers by some of the leading voices in psychology's reform movement that also explicitly relate to psychologists' research on rationality and biases.[99] In the rest of this section, I will try to reconstruct and make explicit their indigenous epistemology from these papers – what is the rationality that the reformers want psychology to be governed by?

The aim of the citation analysis is not to make a strong claim based on scientometrics, but to illustrate and highlight parts of the literature in the reform debates that are relevant for the talk of biases. Neither decision-making psychology nor the reform literature are fully represented. Instead, small parts of those literatures are highlighted in a way that provides an ingress point into a larger body of scholarship.

---

[98] For an up-to-date reconstruction of the main arguments of the rationality wars, see Mercier and Sperber (2017, p. 13-48). For a proposal on how to use this new kind of rationality as an outline for a new epistemological approach, see Bishop and Trout (2005).

[99] The citation network was produced using citation data from Web of Science and the openly available program CitNetExplorer (Van Eck & Waltman, 2014a). See Appendix B for a detailed description of the search query, included data, the software tools used, and the steps in the analysis which produced the visualization in *Figure 5.1.*

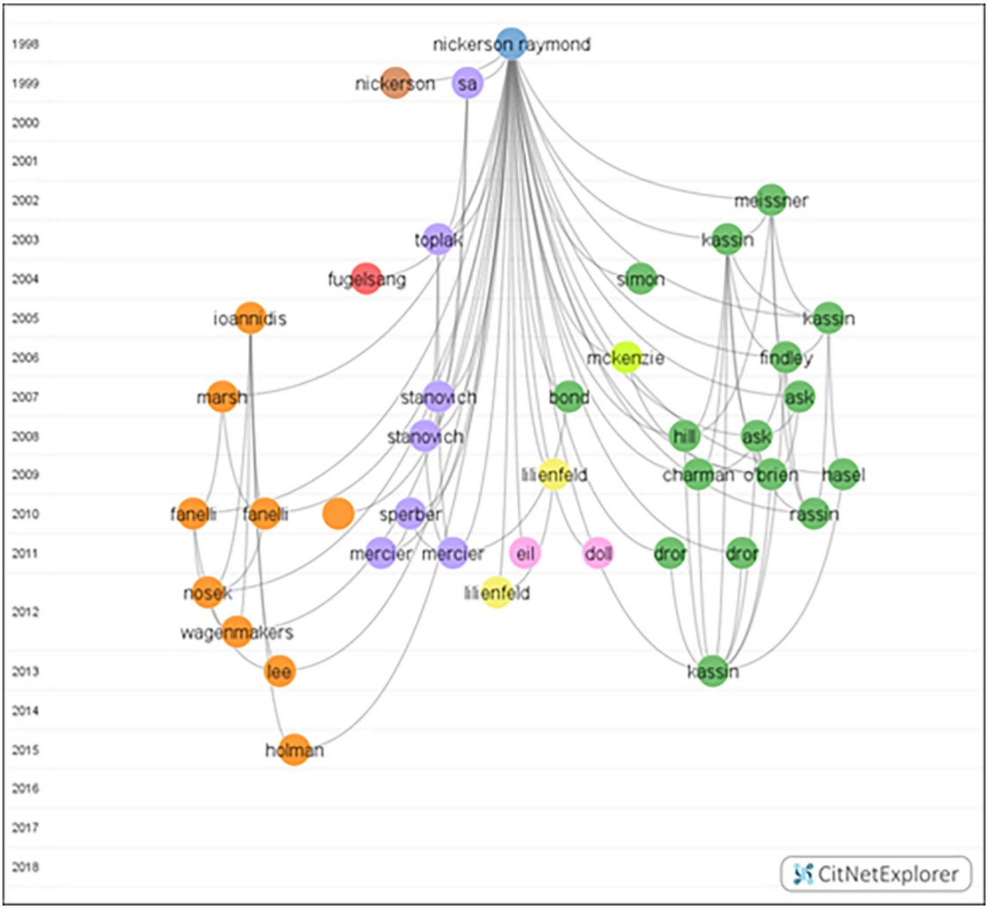


Figure 5.1. Confirmation bias citation network

In the visualization in *Figure 5.1*, some of the highly influential papers[100] published by the reformers directly cite Nickerson's review of confirmation bias: The two papers by Daniele Fanelli in PLOS One (2010a; 2010b); Nosek, Spies, and Motyl's *Scientific Utopia II* (2012); and Wagenmakers and colleagues (2012) *An Agenda for Purely Confirmatory Research*. I will shortly discuss each as representations of the kind of scientist's rationality that the reformers give currency to.

Fanelli published two papers in 2010, *Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data* and *"Positive" Results Increase Down the*

[100] CitNetExplorer visualized 40 publications (out of 1174) in *Figure 5.1*, so the reformers orange cluster is not an exhaustive representation, just the most cited cream of the crop. The clusters of other colors are groups of articles that discuss confirmation bias in various different research lines unrelated to the reform movement in psychology.

Hierarchy of the Sciences. Fanelli's papers are exploring the question of bias toward publishing "positive" results i.e. the results that confirm the tested hypothesis. This is not an individual bias per se, but a kind of bias that manifests itself on the level of published literature. Some scientific literatures exhibit a much higher proportion of positive results than one would expect. In the first paper, Fanelli sets up his research question in the following way (2010a, p. 2):

> [I]f the HoS [hierarchy of science] hypothesis is correct, scientists in harder fields should accept more readily any result their experiments yield, while those in softer fields should have more freedom to choose which theories and hypotheses to test and how to analyze and interpret their own and their colleagues' results. This freedom should increase their chances to "find" in the data what they believe to be true [...],which leads to the prediction that papers will report more negative results in the harder sciences than in the softer.

Fanelli was checking for the manifestations of Nickerson's confirmation bias as a proxy indicator for the question if there is such a thing as a hierarchy of science.[101] Confirmation bias can influence literatures systematically if institutional controls are less successful, and Fanelli makes the argument that this is the case for scientific fields lower in the hierarchy (2010a, Fanelli, p. 2):

> Scientists, like all other human beings, have an innate tendency to confirm their expectations and the hypotheses they test. This confirmation bias, which operates largely at the subconscious level, can affect the collection, analysis, interpretation and publication of data, and thus contribute to the excess of positive results that has been observed in many fields. In theory, application of the scientific method should prevent these biases in all research. In practice, however, in fields where theories and methodologies are more flexible and open to interpretation, bias is expected to be higher.

The basic idea about the psychology of the scientist was spelled out directly here: Science as a system is a set of checks and balances for the biased thinking of human beings. In the more successful scientific fields – the harder ones – these checks are more robust and more thoroughly implemented. The second paper uses the same notion of the biased scientist, and explores it against the backdrop of the geographical distribution of scientific production in the United States. The idea behind this study is that in the states exhibiting more competitive academic environments, the pressure to publish will be greater, thus leading to a more expressed bias toward publishing

---

[101] In the introduction, the paper discusses the Comtean idea of a hierarchy of science through the vicissitudes of many perspectives on science that followed after Comte. Fanelli briefly discusses Comte, philosophers and sociologists of the 20th century, Kuhn and Latour, and concludes with the extreme "postmodern" perspective in which "all the empirical measures of hardness listed above could be re-interpreted as just reflecting cultural differences between 'academic tribes'" (Fanelli, 2010a, p. 2).

positive results. Here too, one of the formative causes that give rise to the positive bias in scientific literatures is the psychology of the scientist, for which Fanelli cites Nickerson:

> Many factors contribute to this publication bias against negative results, which is rooted in the psychology and sociology of science. Like all human beings, scientists are confirmation-biased (i.e. tend to select information that supports their hypotheses about the world), and they are far from indifferent to the outcome of their own research: positive results make them happy and negative ones disappointed. (2010b, p. 1)

An important point that needs to be stressed here is that Fanelli's research does include psychology, but is not primarily focused on psychology as a scientific discipline. His perspective is broader, including science as a whole. This is in line with the newly developing field of meta-science (Ioannidis, Fanelli, Dunne, & Goodman, 2015), or "science of science", spearheaded by the Meta-Research Innovation Center at Stanford University (METRICS). Finding these papers mixed in psychologist-reformers' publications citing Nickerson is no surprise, considering the reform in psychology is an expression and functional part of wider trends in scientific research on science, the Open Science movement, and science reform.[102]

Coming back to the core group of reformers, Nosek, Spies, and Motyl's (2012) paper is one of the two Brian Nosek (the director of the Center for Open Science) published on scientific utopias. In *Scientific Utopia I* (2012), the authors criticize the current state of scientific communication – the way that academic journals and peer review operate and the role they play in the scientific system. The authors wrote their criticism under the title Scientific Utopia "in recognition that we present an idealized view" (p. 218). They go on to argue that:

> The ideas illustrate inefficiencies in the present, and point toward possibilities for improving on those inefficiencies. Although an ideal state is not attainable, it can be the basis for [sic] of improving current reality. Our purpose is to provide a practical template for improving the current model. We argue that the barriers to these improvements are not technical or financial; they are social. The barriers to change are a combination of inertia, uncertainty about alternative publication models, and the existence of groups invested in the inefficiencies of the present system. Our ultimate goal is to

[102] The connection between the meta-science perspective, Open Science, and reform in psychology is manifold – it includes a veritable network of individual researchers, research practices, rhetorical strategies, advocacy, empirical findings, models of the individual psychology of scientists, and sociological and philosophical reconstructions of science. It is also conducive to a sociological (and in the future, historical) "follow the money" analysis (Andersen, Bek-Thomsen, & Kjærgaard, 2012) of the role that funding agencies play in scientific research, considering both METRICS and the Center for Open Science were started with generous funding from the Laura and John Arnold Foundation.

> improve research efficiency by bringing scientific communication practices closer to scientific values.

The plea for reform of the social practices of psychological scientists is even more explicit in the second article on scientific utopias. In *Scientific Utopia II*, the one that does appear in our citation network, the authors discuss the other side of the problem with scientific communication: the incentive structure of science. The current system of incentives, pushing for scientists to publish above else (and in the most prestigious journals they can), just entrenches and facilitates the negative aspects of "ordinary human motivations and biases" (2012, abstract). As the authors put it themselves (p. 616):

> On its own, the fact that publishing is essential to success is just a fact of the trade. Running faster defines better sprinters; conducting more high-impact research defines better scientists. The research must be published to have impact. And yet, publishing is also the basis of a conflict of interest between personal interests and the objective of knowledge accumulation. The reason? *Published* and *true* are not synonyms. To the extent that publishing itself is rewarded, then it is in scientists' personal interests to publish, regardless of whether the published findings are true [...]

Considering that the main incentive actually mixes up the categories of 'published' and 'true', it institutionalizes motivated reasoning. On the surface, it connects "what is good for scientists and what is good for science" (p. 616), while in actual practice it opens the doors for intended and unintended motivated reasoning that gets enshrined and institutionalized. Motivated reasoning lets one publish more so it gets scientists more papers, and by extension, increases their chances to be hired, receive grants, and to occupy gatekeeping positions like refereeing in journals, funding agencies, and hiring committees. Where does Nickerson's confirmation bias come into this? It is used as an example for conceptual replication as one of the "strategies that are not sufficient to stop the proliferation of false results" in the literature (p. 619):

> Because features of the original design are changed deliberately, conceptual replication is used only to confirm (and abstract) the original result, not to disconfirm it. A successful conceptual replication is used as evidence for the original result; a failed conceptual replication is dismissed as not testing the original phenomenon (Braude, 1979). As such, using conceptual replication as a replacement for direct replication is the scientific embodiment of confirmation bias (Nickerson, 1998).

Here, we again see the interaction between an element of the scientist's psychology (confirmation bias) and a social practice of scientists (conceptual replication). The reformers' worry is that the institutionalized social practice, in this case conceptual replication, amplifies and extends the bias into the literature. Nosek and colleagues argue without mincing words that conceptual replication is a form of propagation of biased thinking. Conceptual replication is one of the contested grounds in the reform debates, especially in the case of experimental social psychology. Many of the reformers argue, like Nosek and colleagues above, that the psychologists' criteria for allowing phenomena into the literature were too lax because direct replication was not emphasized enough. As LeBel and colleagues (2017, p. 255) put it: "Social psychology prior to the "replication crisis" (pre-2011) focused primarily on testing generalizability and internal validity via methodologically dissimilar replications, or what have sometimes been called conceptual replications, with much less attention paid to methodologically similar replications, or what have sometimes been called direct replications." Two aims of the reform are, then, first explicitly identifying "the replication continuum" from more to less direct replication, and second, to put more stress on the direct replication end of that continuum in actual research practice. More emphasis on direct replication is not only compatible with the goal of reduction of bias, but it also makes psychological research conform more directly to the Popperian ideal of falsificationism. For falsificationists among the reformers, robust phenomena that are directly replicated provide a pool of concepts ('constructs') on which more dissimilar replications can be performed to establish the construct's conceptual scope (LeBel et al, 2017, p. 256).[103]

The last paper I will use as a representation of the psychology of the scientist at work among the reformers tackles the problem of exploratory and confirmatory research. The confirmatory vs. explanatory research debate very nicely illustrates how the initial choice of phenomena and their conceptual analysis usually take the backseat when it comes to psychologists, regardless of their involvement in the reform attempt. In the now classic paper, Wagenmakers and colleagues (2012) argue for the benefits of confirmatory research in psychology and how preregistration might help (the same preregistration being advocated by Chris Chambers in *Seven Deadly Sins of Psychology*). The paper's opening paragraphs show how, for the reformers, the biased psychology of the scientist is being amplified by accepted practices within psychology as a science (2012, p. 632):

> Psychology is a challenging discipline. Empirical data are noisy, formal theory is scarce, and the processes of interest (e.g. attention, jealousy, loss aversion) cannot be observed directly. Nevertheless, psychologists have managed to generate many key insights about human cognition and behavior. For instance, research has shown that people tend to seek confirmation rather than disconfirmation of their beliefs – a phenomenon known as confirmation bias (Nickerson, 1998). Confirmation bias operates in at least three ways. First, ambiguous information is readily interpreted to be consistent with one's prior beliefs; second, people tend to search for information that confirms rather

[103] The philosopher Uljana Feest (2018, February) makes a similar argument that the current issues subsumed under the name 'replication crisis' need to be separated into the statistical issues and the problems of 'conceptual scope.' According to Feest, a single-minded focus on replication distracts from issues of 'conceptual scope' of psychologists' constructs.

> than disconfirms their preferred hypothesis; third, people more easily remember information that supports their position. [...] In light of these and other biases it would be naïve to believe that, without special protective measures, the scientific research process is somehow exempt from the systematic imperfections of the human mind.

The proposal is to reform the research report in such a way that it will go through a two-step process – the researchers will submit a protocol describing the research design and analysis plan to a journal, and only after this protocol has been reviewed and accepted, they will collect the data and write up the results of the study. In this way, it will be made clear on the level of the report's format what were the hypotheses the scientists started with, and how were those hypotheses investigated and updated through data collection and interpretation. The way scientists in psychology behaved up to the 2010s (and, for the most part, still do) was to conduct exploratory research but write it up and interpret it as confirmatory, without clear separation of the two. The data used to generate hypotheses was the same that was used to test them, packaged into one report that hid the back and forth between data and the researchers' hypotheses. Through registered reports, the way that papers are published is much more closely matched both to the requirements of the inferential tests psychologists usually use *and* the falsificationist scheme for exposing hypotheses to threatening data. Justification of results that prove or disprove phenomena is thus reformed, while the exploratory kind of research remains unreformed and thus allows for "scientific creativity."[104]

The exploratory vs. confirmatory research debate seems to be running into similar issues as the previously mentioned discussion about conceptual vs. direct replication. Both can be mapped onto the philosophers' distinction between the context of discovery and context of justification. However, I don't think that seeing either of those debates as a restatement of the context of discovery/justification will solve them, or even point towards the direction of a satisfactory solution. What it does reveal is something deeply troubling – that psychologists' institutionalized tools and practices for justifying findings are complex and well developed, but the ways for initial selection and identification of phenomena of interest are, in comparison, underdeveloped. Psychologists, especially social psychologists, seem to lack a way for collaboratively producing a wider consensus on what should be the concepts of interest in the first place. Paradoxically, in practice, they can ignore this and go on behaving like good falsificationists with the concepts that they do end up studying. Put crassly, the conventions of psychological research are elaborate when it comes to

[104] I mention scientific creativity here because one of the common worries psychologists express when it comes to registered reports is that they will "prevent exploration of data and curb scientific creativity" (Chambers, 2017, p. 185). The reformers usual answer is that they are just calling for clear labeling of what is confirmatory and what is exploratory in the literature. Psychologists are free to conduct "exploratory" research, as long as they are clear that that is what they are doing.

the rules for justifying findings, but they perform much worse when it comes to mustering discoveries worth justifying in the first place. The elaborate and statistically sophisticated discussions about justification of psychologists' claims obfuscate the lack of capacity for organized discussion about what should be justified in the first place.

To sum up, my aim in this section was to show that the indigenous epistemology at work in the current debates is that of the biased mind of the scientist. The reformers employ a concept of rationality of humans-as-scientists that has ontological implications – scientists think and act in a biased way *because* they are humans, and the science system should be set up in such a way that it complements human (ir)rationality. The cause for reform is the way scientific psychology works on the level of institutionalization of research and publishing practices. With the current institutionalization, the science of psychology not only fails at correcting those biases, but it amplifies and reifies them as peer reviewed literature. For the reformers, humans are fundamentally irrational, so much so, that this threatens the very functioning of science.

5.3 Science is/should be governed by rationality! What rationality and what science though?

5.3.1 What rationality?

The indigenous epistemology taking form in the reform debates of the 2000s and 2010s is that of the irrationality of scientists. This indigenous epistemology was taken as a model of human reason from the rationality wars of the 1990s. The "model" is not formally explicated for the most part, but it spells out a few important elements. The first one is that humans do not conform to rules of formal logic and probability theory when making inferences. Since psychologists-scientists are humans, they are prone to the same biases. Among different psychologists discussing human rationality, 'not conforming' means different things. Some provide normative interpretations, of the kind that say that the discrepancy between human thinking and logic/probability theory means that humans are irrational; while the less normative ones just state that human thinking is well-adapted for thinking about different kinds of problems than those traditionally framed by 20th century logic and probability theory. For psychology's reformers, this distinction is extremely relevant, whether acknowledged or not, because they see any deviation from logic and probability theory as a threat to scientific rationality. However, the second element of the indigenous epistemology tells us that science is not only an exercise in individual reasoning, it is a community enterprise. Many biased psychologists-scientists communicate to make inferences and construct reasonable arguments as a community (or even, a number of loosely related communities). The problem, spurring the reformers to reform, is that the rules and norms of that social construction of knowledge ('reasonable arguments') are broken in two related ways: They allow psychologists to apply formal logic and probability theory in an unsound way and, in the more extreme version, may even incentivize psychologists to do so.

There are two fundamental insights of the reformers' arguments that need to be stressed here: Their belief that (1) scientific thinking is a set of rules for the application of formal logic and probability theory but (2) that application is socially mediated and instituted. Thomas Sturm, in his analysis of the rationality wars, calls this "a fundamentally correct and important insight" of the "bounded rationality" approach to human inference in general: "[B]ecause reasoning often has to proceed on the basis of very little information and large amounts of uncertainty, it makes little sense to expect logic or probability theory alone to be sufficient in a comprehensive normative theory of rationality" (2012, p. 78). The reformers are only indirectly after a normative theory of rationality – they need one to reform psychology in such a way that it would make psychologists doing research conform to its norms. For the reformers, the normative theory of rationality is not an object of research, but a tool for reforming the discipline. And here, I think, there is room to criticize the reformers because they complement their indigenous epistemology of irrationality with an outdated model of science as a system.

At this point, a clarification is in order regarding "bounded rationality." The view that rationality is bounded, in the sense of Herbert Simon and Gerd Gigerenzer, mean that problem-solving (more broadly, and in particular, in science) is bounded by the uncertainty of the environment, limitations of human cognition, and the finite time at the disposal of the problem-solver. In such a view, using heuristics is a necessary strategy because of the interaction between humans' cognitive limitations, the uncertain environment, and the time constraint. One cannot reach perfect solutions, rather, one can look for the best possible solutions. What the reformers seem to argue is not only that the science system needs to be redesigned into an environment that does not produce uncertainty in itself, but that at the same time acts as a corrective for the cognitive limitations of individual scientists.[105] For science to act as a corrective, it needs to starts acting as a social system of rules and institutions that exhibit the features philosophers of science have identified as crucial components of the scientific method. In other words, the reformers' argument goes as follows: Philosophers have produced a reconstruction that turns science as a social process into an exercise in logic and probability theory, and scientists should work on

[105] "Environment" in this context has two meanings. The first meaning is the focus of the reformers – it consists of the institutions, practices, and norms individual psychologists are constrained with when they reason scientifically. The reformers argue that this kind of an environment produces so much uncertainty or even systematic bias, that the environment in the sense that is actually of interest to psychologists, the one psychological researchers are after which includes the psychological objects of research (attitudes, personality, intelligence, cognitive functions, heuristics, etc.) remains inaccessible. Put simply, the scientific system, in case of disciplines like psychology, started producing so much uncertainty that it makes the already uncertain environment of psychological functioning and behavior inaccessible in principle.

refurbishing the current scientific system along those lines so it could act as a corrective for their own biased thinking. The brunt of the reform falls onto the kind of reconstruction of the science system that psychologists recognize as best, and to that, I will turn next.

5.3.1 What science?

The episode that brought issues of replicability in psychology truly under the spotlight of the wider scientific community was the publication of the large-scale collaborative replication study by the Open Science Collaboration (2015). The humble conclusion of the paper which later lead to some of the most fundamental criticism of 21st century psychological science is a good starting point for looking for reconstructions of the science system that the reformers find persuasive (2015, p. aac4716-7):

> After this intensive effort to reproduce a sample of published psychological findings, how many of the effects have we established are true? Zero. And how many of the effects have we established are false? Zero. Is this a limitation of the project design? No. It is the reality of doing science, even if it is not appreciated in daily practice. Humans desire certainty, and science infrequently provides it.

Their study was a huge endeavor involving dozens of labs, by their own evaluation a vital one for the functioning of science. Because of that, for the reformers, the replication effort is the result of scientists just going about their business: "Scientific progress is a cumulative process of uncertainty reduction that can only succeed if science itself remains the greatest skeptic of its explanatory claims" (2015, p. aac4716-7). The falabilist description of science as "uncertainty reduction" has two distinct overtones – first of the indigenous epistemology of human irrationality I have described in this paper; and secondly, of early 20th century conceptions of science as a collection of empirical and theoretical propositions connected by rules of formal logic and probability theory. In that view, science is a collection of propositions produced and checked by the scientific method. This should not come as a surprise, considering the first six citations of the Open Science Collaboration paper: four to three philosophers of science Carl Hempel (1968; Hempel & Oppenheim,1948), Imre Lakatos (1970), and Wesley Salmon (1999); one to the psychologist-philosopher Paul Meehl (1990a) and one to John Platt's highly influential paper *Strong Inference* from 1964.

The perfect illustration for the view of science that has traction among the reformers comes from Munafò and colleagues' manifesto (2017) in its *Figure 1*. The figure's caption reads as follows (2017, p. 2):

> **Threats to reproducible science.** An idealized version of the hypothetico-deductive model of the scientific method is shown. Various potential threats to this model exist (indicated in red), including lack of replication,

> hypothesizing after the results are known (HARKing), poor study design, low statistical power, analytical flexibility, P-hacking, publication bias and lack of data sharing. Together these will serve to undermine the robustness of published research, and may also impact on the ability of science to self-correct.

Each of these threats endangers one or more steps of the hypothetico-deductive model. The steps presupposed by the model, that are shown in the figure, unfold in the following repeating circle: 1. Generate and specify hypothesis; 2. Design study; 3. Conduct study and collect data; 4. Analyze data and test hypothesis; 5. Interpret results; 6. Publish and/or conduct next experiment. As the arrow between the last step and the first shows, a scientist adds to the network by publishing, and she generates and specifies hypotheses by consulting what is published. Science is described algorithmically, as a type of data production. That data is then interpreted and used as evidence to decrease uncertainty. The vehicle for that interpreted evidence is a journal article.

For the reformers, science as a system of knowledge is a network of empirical statements, theoretical constructs, and operationalizations that connect them. This network is maintained by the scientific method – a consistent set of inductive practices for producing data and making inferences about them. "Cumulative scientific progress" is the ordering, expansion, and checking of this network and the practical (albeit) imperfect proxy for that network is the scientific literature as a whole.

Does the above description of science mean that the reform movement is only the newest attempt to make psychology as a science conform to some logical empiricists' views of the proper functioning of science? Or in other words, is the thing that Laurence Smith persuasively argued *against* in his book about neobehaviorism happening today? I don't think so, because the reformers are a plural group of practicing scientists who don't necessarily belong to the same epistemological club, as far as their philosophy of science goes. Sometimes they cite and discuss Carl Hempel, at other times Karl Popper or Imre Lakatos, then Paul Meehl, Mertonian sociological analyses, methodologists like De Groot; and hero-scientists like Feynman.[106] The reformers use all these reconstructions as a backdrop – as a canvas of possible idealizations they find attractive enough to use for describing scientific practice. They need some sort of a formal description of science as a system for their reform punch to land. The reform requests are informed by their indigenous epistemology of irrationality, but the question what *was* science before it got reformed, and what *will* some future science be after the reform, is much opaquer.

[106] For a much more detailed and critical evaluation of the kind of philosophy of science the reformers take inspiration from, see Derksen (submitted).

In that opaqueness – using a very unsystematic number of examples of what is science from conventional analytical philosophy of science – I see the biggest intellectual and practical weakness of the reform movement. Intellectual weakness because it uses, as a crutch, a thoroughly compromised system of thinking about science. Practical weakness because such a simplistic "logic of science" view is neither persuasive nor efficient. No actual science worked like that, neither physics nor biology nor psychology. Scholarly fields that investigate science today, like philosophy/history/sociology of science, have moved away from the late 19th century and early 20th century conception of science as special because of the scientific methods that scientists use. As Paul Hoyningen-Huene puts it, since the last third of the 20th century, the "belief in the existence of scientific methods" that are specially equipped for producing infallible knowledge "has eroded" (2008, p. 168).

Since there is no consensus over a universal and abstracted system of scientific methods, the reformers are fitting their indigenous epistemology of irrationality to an abstraction that has no real import for psychology as a science. They are rebuilding the rundown castle of their discipline from Chambers' metaphor into a castle in the sky – one that works, but has never existed. More to the point, the reformers' indigenous epistemology of irrationality, as I reconstructed it, is a similar position to W.V.O. Quine's naturalized epistemology, a position thoroughly incompatible with Popper, his logical empiricist predecessors, and all other non-naturalistic epistemologies.[107] The reformers, using their indigenous epistemology of irrationality, are after an epistemological system that prescribes "excellence in reasoning", while the conventional philosophies they use as crutches aim at providing justification of knowledge claims (or belief tokens, as philosophers like to call them).[108] Rhetorically, they seem to profess Popperianism or some brand of logical positivism, but in practice, they are developing a thoroughly psychological naturalized epistemology. If with anything, the reformers' indigenous epistemology might be compatible with Latourian readings of science, or other post-Kuhnian sociological or historical reconstructions that the reformers almost never invoke or use. At least those views see scientific practice as thoroughly social and historical; especially when we take into

[107] Some sort of at least lip service to falsificationism seems to be prevalent in the reform movement. Fulfilling the criteria of falsifiability seems as a straightforward road toward the reduction of bias. The fact that the logic of falsifiability runs into insurmountable problems when one tries to turn it into an applied methodology is rarely discussed, because such practical questions about psychologists' research practice fall squarely into the domain of statistical inference, not philosophical analysis.

[108] A book-length treatment of exactly this distinction is the previously mentioned Michael Bishop and J.D. Trout's (2005) *Epistemology and the Psychology of Human Judgment*. As they argue, epistemological theories that focus on reasoning excellence and justification are in conflict in the long run: "A theory of epistemic excellence will yield normative conclusions about the epistemic quality of a *reasoning strategy*. But reasoning strategies typically produce belief tokens. So whenever a theory of reasoning excellence recommends a particular reasoning strategy for tackling a particular problem, it normally recommends a belief token, but at one remove. And this leaves open the possibility of conflict" (p. 17).

account that the kind of naturalism discussed in this paper was actually a psychological naturalism – focusing on observable features of human psychology and behavior that have import for knowledge production. When the naturalistic position is expanded in a more "ecumenical" way (Fuller, 1988, p. 19), including the empirical accounts of science of sociologists and historians of science, suddenly the reformers' take on what science is/ought to be has the potential of introducing a reformed kind of "psychologism" that could be productive both for normative and descriptive accounts of the meta-disciplines like sociology of science, history of science, and STS.[109]

The peculiarity of all the reformers' potentially incompatible positions being mustered side by side, enlisted in arguments they were not really suited for, just goes to show how destabilizing the whole discussion around the replication crisis is for psychologists. Practicing scientists go looking for philosophers' prescriptions when the earth is shaking. I would argue that there's another reason why such a plurality of potentially incompatible philosophical reconstructions of science can coexist in the same reform movement, and that has to do with psychology as a discipline in the late 20th century.

Psychologists since World War II have settled on a methodological standardization of their discipline (Danziger, 1990; Chapter 4 of this thesis). The disciplinary consensus on the use of methods and inferential statistics to use has kept the burgeoning discipline together, but has also neutralized most fundamental discussions about the nature of psychological research, state of psychological theories, and the internal consistency of psychological science as a whole. These discussions did happen in the period since World War II (for example, the one centered on null hypotheses significance testing), but were constrained to specialist discussions of methods and statistics. Squarely atheoretical and anti-metaphysical, scientific psychology was secure in its methodological identity. Even more so because the methods were refined and increased in sophistication and skill requirements – the rise of structural equation modelling and Bayesian statistics being two great examples. Data could be produced, articles could be written, and careers could be made. Literature reviews and meta-analyses would provide a semblance of a structure that promised some future in which the inundation of empirical studies was being integrated into a consistent whole. That secure methodological identity, though, has been destabilized in the 2010s with the beginning of the replication crisis.

[109] Steve Fuller's (1988) social epistemology, as a synthesis of work in history/sociology/philosophy of science and STS, is a type of more-inclusive naturalism. How would the sociologists, historians, and philosophers accommodate for a new and vocal group of scholars that would attempt to flesh out a psychologism in epistemology? Asking the reformers to take a more active role in explicating their epistemological and metaphysical positions is the first step for developing such a productive tension and communication "between" literatures.

Not only is the whole literature that has accumulated potentially untrustworthy, but so are the research practices that have produced it. Psychologists have come to realize that their scientific method is not what they were taught it was in their graduate schools. It was potentially something much more messy and contingent. And that contingent mess, according to the reformers, begs for reform. Late-20th century psychologists' focus on methods is also the reason why, in this chapter, I avoided the elephant in the room thus far – the fact that reformers expend most of their energy in talking about methods and statistics. These are important indeed, but the kind of problems opened up by the replication crisis go deeper, to the fundamental historical and philosophical definitions of the rationality and the science psychologists use to guide and construct those methods.

5.4 Naturalized indigenous epistemologies

The controversial position of naturalism in epistemology is a recurrent theme in this chapter. It is even more controversial in the Western tradition as a whole. At the beginning of the 20th century, Gottlob Frege and Edmund Husserl cautioned against 'psychologism', warning logicians not to muddle their explanations with discussions of human thinking. No wonder it appears again in psychology at the end of the 20th century – the controversy known as the *Psychologismus-Streit* was the cultural and intellectual context that gave rise to German experimental psychology to begin with.[110] Since Wundtian experimental psychology was appropriated into the American traditions of the early 20th century in severely restricted forms, it could serve as a "dephilosophized" model for both the new scientific psychology that developed during the 20th century, and survive to this day as its origin myth. By dephilosophized, I mean that it was shorn from its complex philosophical foundations, which in Wundt's mature phase included an ontologically monist theory of mind, a methodologically plural approach to the subject matter (*Völkerpsychologie* and experimental psychology), and a clearly defined relationship between empirical psychology, empirical psychology's epistemological justifications, and a way for using empirical psychology to build a metaphysics and ultimately Wundt's *Weltanschauung*.[111]

[110] For a thorough sociological treatment over the German debates about psychologism in late 19th and early 20th century, see Kusch (1995).

[111] I base this comment on Saulo de Freitas Araujo's (2016) reinterpretation of the philosophical foundations of Wundt's psychology. The way Wundt's experimental psychology was received during the 20th century might also provide us with a blueprint for how 20th century traditions of scientific psychology first flatten their whole approach in order to remove all of its philosophical implications, thus developing only the methodological aspect of their program. When the methodology flounders, its supporters attempt to save it by extending the methodological view into an indigenous epistemology. Araujo's reappraisal of Wundt's philosophical system might give us another clue for moving forward with the philosophical analysis of 20th century scientific psychology. Wundt developed a theory of knowledge which clearly necessitated his ontological monism *and* methodological dualism (Araujo, 2017, p. 210). All of the psychologists' programs

Later, when psychology was already an established scientific discipline largely disinfected from Wundt-type philosophical foundations, logical empiricists at large constructed a whole system of scientific theories arising from the opposition between synthetic and analytic propositions, and the logical operators connecting them. Between propositions bringing empirical data and propositions expressing logically valid truths, all knowledge could be reconstructed, or so they hoped. By the middle of the 20th century, neither scientists nor philosophers had any use for psychological descriptions of thinking or philosophical reconstructions of metaphysics. Formal logic and empirical facts could dispense with both. That was more or less the case up to the middle of the 20th century. Then Quine's (1951) criticism of logical empiricism, for a short period of time, opened up a venue for a psychologically-informed naturalistic epistemology in the philosophical tradition. In epistemology proper, this window was short-lived, as most epistemologists after Quine thought and still think that "naturalism in epistemology is impossible or self-refuting or self-undermining" (Bishop & Trout, 2005, p. 23.). In the newly formed discipline of analytical philosophy of science in the middle of the 20th century (Reisch, 2005), the criticism of logical empiricism caused a ruckus. Popper's falsificationism was provided as an alternative, then heavily criticized by the likes of Paul Feyerabend and Michael Polanyi, or received extensions in the work of Imre Lakatos and neo-Popperians. Thomas Kuhn's *Structure of Scientific Revolutions* in 1962 signaled and expressed the eclipse of the analytical tradition of describing science, and the rise of sociologists and historians of science who occupied themselves with socially contingent scientific practice as their object of interest. Psychologically-informed naturalism in epistemology, thus, did not survive among philosophers nor among the historians and the sociologists.

However, this kind of epistemological naturalism could survive outside of academic philosophy, history, and sociology of science. The examples of indigenous epistemologies described in this paper show us how for psychologists, naturalism wasn't only a possible position, but a necessary one considering they based their indigenous epistemologies on their theories about human thinking and behavior.[112] For psychologists, science is a product of human psychology. The only question is: What was the current version of the scientific description of the human mind and behavior informing that view?

I've discussed in this chapter seem to be methodologically monist because of the commitment to the methodological standardization since World War II I discussed previously. Maybe a radically new theory of knowledge, akin to Wundt's, is needed in order to relax this methodological standardization?

[112] There is an interesting historical case study to be made here, considering Quine took inspiration from Watson's behaviorism and his productive years overlapped with the height of neobehaviorism. For more, see Quine, (1991). Also see a comment along the same lines by B. F. Skinner (1987, p. 207) in a peculiar review of Laurence Smith's *Behaviorism and Logical Positivism.* I say peculiar because Skinner reviewed a book in which he was one of the most important historical actors.

Experimental psychology, after the 1960s and 1970s, went through the cognitive revolution. The revolution drastically changed some programs of research, influenced others, and by virtue of its multidisciplinarity led some psychologists away from psychology into the newly developed cognitive science. The wider discipline of psychology also expanded enormously, with the application of a stable core of experimental and correlational methods (Chapter 4) serviced by a controversial brand of inferential statistics (Gigerenzer et al, 1990; p. 203-234). In the 2000s, numerous lines of methodological criticism from within psychology and without (the wider 'science of science' perspective and Open Science) started taking explicit form and culminated in the replication crisis. The stable methodological core was exposed to fundamental criticism, and by extension this criticism cast serious doubt on the enormous literature of many communities of psychologists that have been expanding for decades. Experimental social psychology was hit the hardest, but the criticism affects all areas that have internalized the methodological core of late-modern psychology, with its research designs, theories of measurement, and inferential statistics.

In this chapter, three examples of indigenous epistemologies were identified as naturalized epistemological positions that grew out of the research programs of different psychologists during the second part of the 20th century.[113] I would like to explicate a few salient features that make these epistemologies indigenous on the one hand, and naturalized on the other. They are indigenous because they were philosophical formulations about knowledge production that weren't imagined from the outset by their authors as philosophical positions. The comparison with Wundt's project illustrates this nicely: If we follow Araujo's reappraisal of the philosophical foundations of Wundt's psychology, Wundt was developing his empirical psychology with the explicit goal of distilling metaphysically and epistemologically relevant conclusions from it at some point. He was a philosopher developing psychology as an empirical science in order to inform his metaphysics and epistemology. Contrary to that, the indigenous epistemologies of the neobehaviorists, Maslow, and the current reformers aren't devised as self-contained philosophical systems. These psychologists stumbled into the philosophical foundations of their science, they didn't plan for it. The other salient feature of the three indigenous naturalized epistemologies is that they are types of psychological naturalism *because* they are indigenous to scientific psychology. They all give psychological accounts of scientific knowledge production because they were formulated as extensions of psychological research.[114] Here, the

[113] The argument could probably be extended in such a way that psychologists in the first decades of the 20th century also developed naturalized indigenous epistemologies, but that goes beyond this chapter.

[114] This is also the cause of an interesting kind of myopia which sees naturalism in epistemology only as psychological. From the perspective of naturalized indigenous epistemologies of psychologists as I described them, "naturalism" is fully exhausted by giving a psychological explanation of scientific practice. However, if naturalism just implies an empirical attitude toward scientific practice, it also has to include social and historical factors.

similarity between them ends. Maslow was directly opposed to the way neobehaviorists conducted their science. Consequently, he was critical of their views on what is the ideal way to do science. In turn, the current reformers would probably see Malsow's calls for humanizing science as regressive and potentially threatening. My argument is not that all those programs were the same, and that we can easily ignore their differences. I am just trying to call to attention that they share some salient features that should interest us when we investigate the way psychologists describe and *pre*scribe scientific practice.

5.5 Conclusion

I draw the following implications from using indigenous epistemologies as an analytical tool for understanding the reform movement. They are meant as summaries, but also as advice for moving the discussion forward.

If you're an epistemological naturalist, be prepared for a lot of arguments with a lot of people. My impression is that psychologists involved in today's reform debates aren't aware that a psychological naturalistic position on how reason works is highly controversial. It has to be carefully argued and qualified or it will be criticized from almost all communities working on describing science: Sociologists will find it abhorrent because it deemphasizes social context, historians because it potentially essentializes something that has historical contingency; and philosophers because it's in conflict with most of their current epistemological positions. I would argue that this is precisely why the indigenous epistemology of irrationality could get traction among psychologists and meta-science researchers – they are so far removed from all these communities of scholars that they do not share these views as a matter of education. Does that mean that the indigenous epistemology of irrationality is doomed to fail and be forgotten, like Maslow's and the neobehaviorists'? The question whether the indigenous epistemology of irrationality can become viable in the long run is largely tied to two things: a) the precise model of rationality it inherited from the post-Cold War rationality wars and b) the model of functioning of the science system that the reformers couple it with. Both a) and b) necessarily require the reformers to develop an explicit indigenous epistemology that will be consistent with their reconstruction of the science system, before and after reform. They need not become philosophers of science, but they need to conceptually argue for what is scientific psychology without using Popperianism or logical empiricism. In other words, more work needs to be put into specifying what psychology as a science is, extending it beyond lip service to currently popular philosophies of science.[115]

[115] Maarten Derksen (submitted) argues for the opposite point – that the reformers practical implementation of Popper's critical rationalism might lead into all kinds of constructive directions. Even if that is so, it presupposes much innovative conceptual work on what is the agreed upon epistemological position of

a) Which rationality? This is an open question considering the rationality wars have not had a definitive victor for now. However, I do think that the right direction is moving away from pitting 'human biased reasoning' against 'unbiased reasoning of formal logic and probability theory'. A good candidate is Mercier and Sperber's interactionist approach (2017, p. 108): "Reason [...] is a mechanism of intuitive inferences about reasons in which logic plays at best a marginal role. Humans use reasons to justify themselves and to convince others, two activities that play an essential role in their cooperation and communication." For science reform, and psychology in particular, this means that the heavily regulated communication system that was taken for granted by scientists for decades is not a given. Conservatively clinging to a journal system (against peer-review reform and Open Science), article structure (against reform of APA style and critical questioning of reporting styles), methodological prescriptions (e.g. null hypothesis significance testing, but also, more fundamentally, any methodological rule that is taken as a given for decades) needs to be inspected on case by case basis. The same goes for rules of cooperation, or as the reformers call them, incentive structure. This is not to say that I agree with all the points raised by psychology's reformers, just with the approach that sees scientific rationality as a property of a complex social system, and not (only) individual scientists. Consequently, depending what goals scientists set for their activities, that complex social system can function optimally, less optimally, or pathologically.

b) Which model of science? I would like to voice stronger disagreement with the reformers when it comes to their model of science as a system. To put it polemically, taking inspiration from post-positivist or logical positivist philosophy of science makes for strange bedfellows. Instead of looking toward the different articulations of Popper's falsificationism or the hypothetico-deductive model, a much more fruitful source of inspiration for adherents to the indigenous epistemology of irrationality might be the scholarly perspectives that emanated from Kuhn's break with analytical philosophy of science in the 1960s, like historical epistemology and the different schools of sociology of science. As I have mentioned before, psychological naturalistic positions are still nominally[116] incompatible with those reconstructions of science, but at least those scholars speak of science as a complex social system, and not an abstract system of statements serviced by a mythical scientific method. There is another advantage of turning in that direction: There are vital communities of contemporary historians and philosophers of psychology, historians of the human sciences, and critical psychologists whose work might be a more suited source of polemics, ideas,

psychologists and how to put it to work; not only technical and statistical improvement leading to methodological rigor.

[116] I say nominally because some scholars have already developed accounts on how to compare and productively contrast psychologism, sociologism, and historicism in the description of knowledge production. For an STS take on it, see Fuller (1988, p. 3-30). For a philosopher of science's take on it, see Goldman (1999).

and productive intellectual conflict than decrepit analytical philosophy of science. The question of meshing psychological naturalism about science with that of the sociologists and historians is an open field that might be extremely productive not only for normative reform of psychology, but also for descriptive psychology of science. In this paper, I have tried to lead by example: Join the reform debates by discussing work from a newer kind of history and philosophy of science/psychology, like that of Erickson and colleagues (2013), Laurence Smith (1986), Boris Kožnjak (2017), but also other philosophers and historians. In the same vein, I will conclude with the words of Lorraine Daston (2015, p. 676) in her commentary of an *Isis* focus section on history of science and bounded rationality:

> But if confronted with a choice among rationalities, as many philosophers and scientists now believe themselves to be, would it be more rational to prefer knowledge to knowing, efficient procedures over understanding? A history of rationality that took full account of the protean forms packed into that deceptively singular term cannot make that choice, but it could at least illuminate the options and their origins.

In other words, complementing the indigenous epistemology of irrationality with a contingent, messy, historical, social, and plural account of what is science might propel the reform movement in a much more constructive direction than outdated philosophy of science of the mid-century. Who knows, it might even give us a completely new science of psychology!

Chapter 6. Conclusion

This thesis presents the results of an investigation of late twentieth century scientific psychology. I looked at scientific psychology as a type of discourse on human psychology and behavior. Instead of emphasizing the many changes in that discourse, I described the common features in the way that psychologists produce and argue about psychological knowledge. My primary aim was to describe the institutionalized conventions of scientific psychology in the period of the last seven decades. My approach focused on psychologists' discursive practices and their methodology.

The discussion of the discursive – the psychologists' ways of arguing – was manifold. First, on the basic level, my analysis dealt with narration. When looking at textbooks, I based my conclusions on the way psychologists as authors of textbooks narrated psychology to students. Authors of textbooks, in the form and content of the language they used, modelled for students what scientific psychology is. They narrated the history of scientific psychology, the kind of natural scientific methodology preferred by its practitioners, and the unfinished project that psychology was as a scientific domain of knowledge.

Psychologists also instructed the students through the form of the language they used. The most obvious example is in the referencing style, where each piece of evidence is introduced with a year of production and a name of the producer. I also looked at the way psychologists wrote *about* textbooks. Teachers of psychology developed their own historical arguments when writing about the textbooks' past; a type of historical argument separate and distinct from the one used by historians who write about the same books. Teachers of psychology writing about textbooks operationalized them as vehicles of content. 'Textbooks as vehicles of content' are something very different than the instructional manuals Morawski, Smyth, Watson, and I (Chapter 3) tried to reconstruct and investigate.

Lastly, I used a different analysis of written language to retrieve the discourse on scientific psychology when describing psychologists' academic journals. With the academic journals, my focus on the content of psychologists' arguments was comprehensive, but cursory. We data-mined the terms that psychologists most commonly used in the abstracts and titles of the articles they published during a period of fifty years. By data-mining so many titles and abstracts, and focusing on the most frequent terms used by psychologists, we reconstructed a pattern the literature exhibited. This high-level pattern was stable during the whole period. By high-level, I mean it was extremely general (it will arise only when we extract terms out of many articles), cursory (it arises out of titles and abstracts), and abstract (it is abstracted from many of the particular research lines, and it that sense, it is too coarse to be applied specifically).

Seen in that way, both the parts of the thesis looking at textbooks (Chapter 2 & 3) and journals (Chapter 4) retrieved how scientific psychologists' argue about the content of their science. In the textbook case, I described the idiosyncratic way a group of textbook authors narrate psychology (Chapter 3) and the way such narratives are re-described by teachers of psychology on the one hand, and historians on the other (Chapter 2). In the journal case of language use, I looked at what terms and names of concepts were given currency to by thousands of authors writing psychological articles during fifty years (Chapter 4).

By looking at the way psychologists argue about their discipline and about the knowledge claims that are produced within the confines of that discipline, my aim was to provide a description of the institutionalized conventions of scientific psychology in recent history. I provided the basic outline of these conventions in the introduction and Chapter 5. In the introduction, I argued that scientific psychology consists of a number of practices that were turned into bureaucratic conventions. These conventions are the black-boxed investigative practices of 20th century American psychologists. Black-boxing investigative practices means turning them from local, culturally constrained ways of arguing about knowledge into a set of procedures that is perceived by its users as ahistorical, non-local and by virtue of such universality, as objective and necessary. In Chapter 5, my argument was that these conventions turn from implicit to explicit when psychologists start arguing about epistemology in the face of controversy. I described this using Laurence Smith's concept of indigenous epistemology – psychologists' system of thinking about their own science.

When arguing about methodology in the thesis, I used the concept in a narrow sense and a broader sense. In the narrow sense, methodology consist of the ways psychologists design their studies and analyze their data. This is a common intuition among psychologists, that I share by virtue of my undergraduate and graduate training in the discipline. Methodology is the way research studies are designed (experimental, quasi-experimental, correlational, etc.) and the way data from these designs are analyzed (by certain statistical procedures such as correlation, ANOVA, factor analysis, structural equation modelling, or inferential statistics like NHST or Bayesian statistics).

In the broader sense, methodology is the repackaging of the conventions of scientific psychology I outlined in the introduction. The procedural and structured way of thinking prescribed by recent scientific psychology consists of 1) inferential statistics as a way of drawing conclusions, 2) operationism and 3) construct validity theories as a way of structuring the content of research, 4) the literature as an unfinished nomological network of psychological theory, and 5) the genre of writing that connects all those basic elements and reproduces them.

The main empirical conclusions of my analysis are the following. They are called empirical because they are summaries of my work on textbooks and journals:

Empirical conclusion 1: Scientific psychology in recent history (1950-2000), especially in the global north, is sustained by a collection of institutionalized conventions.

Psychological research in this period is extremely varied in content and outlook, but out of that variety, a set of norms and practices can be identified as stereotypes of what scientific psychology should be. It is quantitative, it bases its conclusions on inferential statistics, it defines its content through operationalization of constructs, and it accumulates with the goal of producing a more comprehensive system of knowledge in some, yet unachieved, future. Most subdisciplines of scientific psychology try to conform to these conventions. If they do not conform, but wish to be perceived as scientific, the members of the subdiscipline must invest effort into arguing and justifying their lack of adherence to the particular convention.

This conclusion is, in part, nothing new. Historians of psychology who talk about the history of methods in the 20th century have, in some shape or form, made similar arguments. My addition to that historical work is an elaboration of that view of scientific psychology through the chronological analysis of two discourses: One pedagogical, appearing in published textbooks, and the other data-mined from specialist journals.

Because of the ways I delimited scientific psychology in this thesis, one more aspect becomes visible. Psychologists at large still perceive the institutionalized conventions of scientific psychology as mostly immutable. In their view, scientific conventions are not adaptable tools one can change, adapt, or discard. Rather, they are hard-won givens that psychology cannot do without. In the past decades, however, historians of psychology have persuasively argued that psychologists' investigative practices are part and parcel of the contingent historical process. Scientific descriptions of the human mind and behavior (and the methods for the production of those descriptions) are always embedded, and that embedding is not trivial. Their connection to the wider cultural, societal, and intellectual worldviews is not a methodological issue to be resolved, but an inescapable feature of humans arguing about human nature.[117] While a truism for historians, this insight has left psychologists' research practices largely untouched. For scientific psychologists, "objective" and "scientific" still means "context-independent on the level of tools used to describe things of interest." For historians, the insulation of methods and knowledge claims is an end-point of a successful scientific tradition, not its defining feature or even an indicator of

[117] I use human nature here instead of the much less loaded "human mind and behavior." Human mind and behavior is much more appropriate for recent history, but in a true *longue durée* view, scientific and philosophical views on human mind and behavior would be subsumed under the historically variegated conceptions of human nature (Smith, 2007).

verisimilitude.[118] Put simply, successful scientific traditions are not successful *because* they seem ahistorical and decontextualized, but they seem to be ahistorical and decontextualized because they are successful.

This distinction between historians of psychology and psychologists is a relevant point regardless of the metaphysical position we take. The argument is *not* that all historians of psychology are social constructivists, while scientific psychologists are not. Neither do I claim that some kind of social constructivism is the only correct metaphysical position. Saying that historians of psychology (and more broadly, historians of science) are by definition social constructivists is wrong because many are not.[119] Scientific psychologists also exhibit a whole gamut of positions, or, for the most part, do not explicitly state their views on epistemology or metaphysics because it is largely irrelevant for their particular research practice. Discussions of metaphysical positions or epistemological views are not relevant to my aim either, which is to discuss the most productive match between scientific practices and phenomena. I follow those historians of science who contend that social constructivism does not automatically imply relativism. Investigative practices can be both socially constructed and real, as Daston (2009, p. 813) puts it:

> Probably most historians of science these days, if asked about an episode like the refinement of precision measurement techniques or the formulation of statistical correlations, would answer that such scientific practices are both socially constructed *and* real. That is, they depend crucially on the cultural resources at hand in a given context (mid-nineteenth-century industrializing Prussia, early twentieth-century eugenics-obsessed Britain) and they capture some aspect of the world; they work. But they are neither historically inevitable nor metaphysically true. Rather, they are contingent to a certain time and place yet valid for certain purposes.

What happened in scientific psychology in the past seventy years, I argue, was that the productive traffic between the "given context" and the "ability to capture some aspect of the world" was mediated through highly solidified conventions for thinking and researching. Scientific psychologists have hedged their bets from mid-twentieth century onwards on a set of methodological prescriptions. The *sine qua non* of those methodological prescriptions is that scientific psychology's "ability to capture some aspect" of the human mind and behavior is *made possible through, and only through, isolation from a "given context."* The isolation is achieved through operationalization, constructs, inferential and descriptive statistics, a structured genre of writing, and literatures as nomological networks of constructs. The issue psychologists run into

[118] As Bouterse puts it "the entities that constitute science and the entities that constitute nature (as the object of scientific interest) are in continuous interaction with each other. There is no meaningful sense in which we can claim there to be a 'net influence' of society, or of nature" (2016, p. 183).

[119] Daston (2009) elaborates the distinction between social constructivism and historicism.

here is that the very object of research they want to describe is the crossroad between nature (20th century brain sciences), our conscious set of meanings and labels and names of things (psychological sciences, like scientific psychology or psychoanalysis), and the rules and norms we work with (the object of social sciences proper like sociology). Gerd Gigerenzer called this tools-to-theories, and indeed, it seems that for psychologists, the connection between their investigative practices and the objects investigated by those practices is closer. Very often they seem to constitute each other. Because of this mutual dependence of object and method, in my view, conventions of scientific psychology stultify constructive thinking about phenomena of interest. If method defines the possible realizations of the object of research, and the method is ossified beyond the possibility of development by practicing researchers, the ability to describe and constitute scientific objects is severely limited.[120] This leads to my second empirical conclusion.

Empirical conclusion 2: Scientific psychology, as a collection of institutionalized conventions in the global north, has intellectual inertia.

Sophisticated histories and overviews of psychology stress how diverse scientific thinking about human psychology and behavior has been, especially in the 20th century. My argument is that in this multiplicity of psychologies, the momentous inertia of scientific psychology as a set of procedurally defined investigative practices and rules for thinking about human psychology and behavior was so deemphasized it became virtually invisible. In the thesis, the inertia is manifest in the stability of the scientific literature, but it is sustained through scientific practices.[121] It was sustained through complicated statistics, research designs, operationism, construct validity's bureaucratized rules of arguing about theoretical concepts, and a highly formalized writing genre. The most visible product of the psychologists' maintenance-through-practice is a virtually unsurveyable literature. Scientific psychology wasn't elaborated in a vacuum, but by the work of thousands of scientific psychologists applying their way of thinking to a dizzying number of topics they wrote about in their subdisciplinary communities. Again, I would like to stress that I am not arguing that all psychologists doing scientific psychology were of the same mind during the last seventy years. Rather, my argument is that their work sustained a collection of

---

[120] This is also why it seems like the only researchers who have the ability to innovate theoretically, by the end of the 20th century, seem to be the 'methodologists'. They seem to have the necessary training and agreed upon language that allows for methodological innovation not accessible to most other scientific psychologists. I will discuss this later on the example of the network view of psychopathology.

[121] Conventions of scientific psychology, the literature, and scientific practices of psychologists, in my account, all seem to collapse into "inertia". This is an artifact of my approach in the thesis – I look at prolonged patterns in the end-products of psychologists' practices (introductions in textbooks, abstract/titles of articles, debates about epistemology among reformers of psychology). The conventions, the literature, and the practices form feedback loops, which I may have delimited too strongly in order to be able to describe them.

institutionalized conventions of scientific psychology that has normative and practical import on how psychological research is done.

These two empirical conclusions lead to the critical conclusion of this thesis:

Critical conclusion: Psychologists, through the widespread application of scientific psychology, have enforced norms on arguing about psychology that do not allow for constructive conceptual analysis.

Psychologists working within recent scientific psychology have built a discipline through method. The rules for their thinking about phenomena and their scientific collaboration for investigating them are set. Scientific psychologists are collectively adding to the incremental descriptions of those phenomena. The sheer size of their communities and the way those communities access and add to a seemingly organized literature through curated collections of journals (whether in databases or review/abstracting journals), creates the image of a comprehensive and organized body of knowledge. My argument is that this is not really the case, because the work done on that body of knowledge lacks an organized conceptual framework.

The inertia of scientific psychology has had a stunting effect on the way psychologists argue about psychological knowledge because of the combination of two factors. First, it definitely did allow for an unprecedented productivity of scientific psychologists. Researchers following the conventions of scientific psychology can produce enormous amounts of empirical research. They have a system of data production and interpretation that is on a par with the pace of production of natural science research groups. Second, this cornucopia of empirical research has a big downside. The scientific psychologists' ability to organize and analyze that flood of empirical studies is disconcertingly limited. Scientific psychologists do conceptual analysis by discussing how they operationalize constructs, how those constructs interact with each other, and which inferences are warranted based on the data collected. In this way, conceptual analysis of psychological phenomena is seriously limited. A psychologist can talk of samples, constructs, operationalizations, but not much more.

The limits of this conceptual analysis are highlighted in the replication crisis debates. As I have argued in Chapter 5, the replication crisis exposes the fault lines of scientific psychology. A rigid system for the production of empirical studies does not equip the producers for a collaborative way of deciding what is the nature of the abstractions (theories/models) they are trying to infer from their data. It is also at a loss when it comes to specifying what are the epistemic goals of psychologists' research. Is the aim of psychologists' research to produce internally consistent and comprehensive theories? Or to fit data to models? Or to produce disconnected sets of evidence about various phenomena of interest? All three, or none of the above? If one or all three, according to what criteria do psychologists decide when theories are sound (or when the models fit the data well)? Keep in mind that this is a recurrent problem for all scientists, not just psychologists. However, a complication specific to the communities of twentieth century scientific psychologists is that they only generate methodological solutions to this problem.

When the "methods talk" breaks down, as is the case with the deadlock of experimenters' regress that the participants of the replication debates ended up in, psychologists turn to indigenous epistemologies and philosophy of science. Their lack of vocabulary and disciplinary habits for arguments about the epistemic goals of scientific psychology pushes psychologists to look elsewhere. And the elsewhere, that should hold enough epistemic authority recognized by psychologists, is either their own currently dominant theories or the currently influential philosophy of science. There is a peculiar circularity in the way that psychologists use their own current theories to explain the issues at stake in the replication crisis. If psychology's theories are epistemologically suspect, can they be used to clarify and fix the institutionalized problems of scientific psychology? The circularity is usually easily ignored, because not all psychological models and theories are criticized. Some research lines seem to be safe from fundamental criticism. However, the lack of consensus over how far the replication problems reach into psychology's literature leaves this problem of circularity open. Could it be that even decision-making psychology is built on shaky foundations? Only time will tell, but for now, many reformers use it as a tool for fixing psychology. In Chapter 5, I have described both the turn to psychologists' own theories and philosophy of science through the analysis of indigenous epistemologies.

Firstly, psychologists, when faced with the need for conceptual analysis, turn the issue of setting epistemic goals into problems of their own research practice. They naturalize their knowledge production. Neobehaviorists conducted conceptual analysis by framing the epistemic goals as goals of behavior. Abraham Maslow performed a similar move, reframing the analysis of epistemic goals as the analysis of scientists' personality and the way scientists achieve self-actualization which allows for creative work. The reformers in the 2010s reverted to decision-making psychology, conducting conceptual analysis through the description of the seemingly newly discovered problem of the interaction between scientific norms and the biases of individual scientists.

Secondly, and especially when pressured with a crisis of confidence in their science, psychologists turn to philosophers of science for guidance in conceptual analysis. The various episodes of the interaction between logical empiricism and scientific psychology in the 20th century provide many historical examples of this. In this thesis, I have discussed at length Laurence Smith's analysis of the interaction between neobehaviorism and logical positivism, but also other interactions, like the role of Stevens' operationism which had an impact on scientific psychology in a more complex way than just the old story of behaviorists embracing positivism. Another common recourse to philosophy of science is the resurfacing of the hypothetico-deductive model and the view of scientific literatures as nomological networks in the 2010s reform debate. The use of neo-Popperianism as the mode of conceptual analysis is also quite widespread.

My argument is that neither naturalized epistemologies nor current philosophies of science can substitute actual conceptual analysis in psychological research. If my assessment of institutionalized conventions of scientific psychology and their spread is even partly correct, the majority of psychologists are continuously producing research and building it into a literature that is not adding up in the way that they think it is. And they have been doing so for generations. If that is so, then the scientific production of psychological facts is just that – fact-production, and I take it that few psychologists will consider this to be enough. Consequently, I think two more practical conclusions can be drawn from my research on the inertia of scientific psychology. I wish to conclude with them, because my intended audience in writing this thesis are not just historians and philosophers of science, but psychologists doing research in their different specializations. I do not wish for my “work that is critical of science” to stay “hidden away in places where scientists are unlikely to see it” (Brock, 1995, p. 29).

Psychologists should look for productive metaphors in different places. For most of the 20th century, scientific psychologists have looked upward the perceived hierarchy of science for productive metaphors. From operationism, to the different philosophies of science developed by analytical philosophers usually talking about physics, to the very idea of psychology as a natural science. Throughout the chapters of this thesis, I have argued on numerous occasions that historical thinking might be a different, more productive source of inspiration. This argument could be expanded even further, to Thomas Teo’s (2017) suggestion of psychological humanities as a kind of lively conduit of concepts and ideas between scientific psychologists and the disciplines on that other side of C. P. Snow’s (1959) two cultures. History is just one of them. This is not to say that scientific psychology should become a humanities discipline, but that the humanities’ vastly different ways of talking about human psychology are a productive source of concepts, metaphors, and theoretical structures. Being reflexive, conceptually sophisticated, and oriented toward language, historicity, description, or meaning does not mean being unscientific. Conversely, the biggest source of scientistic anxiety for psychologists are scientific psychologists themselves.

I think historians of science are a great model community for a more productive kind of a hybrid. Well-versed in scientific thinking, historians of science still remain scholars oriented toward interpretation and hermeneutic understanding of their subject matter.[122] Take my previous example of Lorraine Daston drawing the nuanced distinction between relativism and social constructivism when talking about her object of interest, scientific practices. Could psychologists talking about personality, intelligence, mental disorders, social inequality, and heuristics draw such a distinction between psychologism and social constructivism or realism? What would such a

[122] See Jeroen Bouterse’s (2016) work on the hermeneutic philosophy of historiography of science for an in-depth exploration of the relationship between scientific descriptions of the natural world and the historians describing them.

concept of personality or mental disorder even mean?[123] These questions are barely comprehensible within scientific psychology. What is psychologism or a psychological level of description and explanation in the first place? Among psychologists the social constructivists, realists, and agnostics can coexist without communicating while publishing on the same concepts ("constructs") within their separate literatures. As long as your research falls under the conventions, you can commit to your concept of choice in whatever way. Constructs, operationalizations, and inferential statistics neuter most fundamental disagreements, and you can end any article with the invocation of "more research is needed on this topic". This leads me to my last conclusion on conceptual work.

Discussing methods is not conceptual work. Discussing research methods, sampling, and statistical analysis is not the end-all of conceptual work. Twentieth century psychologists have a long track record of ignoring the oppositions, sweeping the difficult questions under the rug and forgetting about them. Gerd Gigerenzer (1998) calls this kind of avoidance of conceptual work by psychologists as producing "surrogates for theories" – a kind of activity that looks like theoretical work, but it actually is not. In this thesis, I have proposed a number of questions where such conceptual work is desperately needed. I will explicate a few, with the hope that it will motivate psychologists to come up with creative answers: What is psychological theory? What is the relationship between prediction, explanation, and theory? Is psychology's literature a proxy for comprehensive theory? Is the accumulation of research studies an optimal way of producing psychological knowledge, or does this accumulation have a different goal? What is that goal? What does it mean to replicate studies, and what does that tell us about constructs/phenomena/effects? Is scientific psychology looking for universals? If it is not, what is the aim of psychological descriptions of phenomena?

Many of the above questions were directed at specifying what psychological theory actually is. The answers might go in a wholly different direction, a direction away from theory. Maybe psychologists should drop the idea of talking about psychological theory. Maybe the intended products of psychological research are huge systematic collections of evidence about particular phenomena; not systematized theory about

[123] Here, there is a struggle for invoking already existing examples of good practice. One instructive example might be the newly developed network view on human psychopathology, which sees mental disorders as feedback loops of symptoms instead of diseases (Borsboom, 2017). The network view is, especially in the way that it connects different levels of explanations, highly reminiscent of Hacking's (1996) "looping effects on human kinds." I mention the network approach to psychopathology here because it shows how a change in the usual methodological tools at the psychologists' disposal yields huge potential gains in reconceptualizing phenomena. The biggest gain, hinted at by Borsboom (p. 11), might be in that the network approach in psychopathology can act as an "organizing framework". Mental disorders as symptom networks are a good outlook for combining mathematical formalism, phenomenological experience of individuals, and the varied relevant levels of explanation. It remains to be seen whether it solves the conceptual disarray most scientific psychology seems to have found itself in.

those phenomena.[124] My point is that whatever the agreed upon goal of producing psychological knowledge is, it needs conceptual work that is today almost never done by psychologists in an organized, institutionalized, collaborative, and funded way. Conceptual work is needed both in setting the goals of psychological research, and in the way the phenomena under investigation are discussed. Scientific psychologists just produce studies, come what may. A mass of studies might have produced a diverse literature, but it has done little in the way of producing cumulative knowledge.

[124] For one way of doing this, see the project of the behavioral change ontology in health psychology (Larsen et al., 2017). An ontology, as elaborated in information science, "is a systematic method for carefully articulating the inter-relationships between classes of carefully defined 'things' or phenomena we care about" (Larsen et al., 2017, p. 7). A good way of thinking about it is a computationally intensive way for solving the conceptual problems in psychologists' talk of constructs that Kathleen Slaney (2017) described.

References

Abma, R. (2013). *De publicatiefabriek: Over de betekenis van de affaire-Stapel*. Nijmegen, The Netherlands: Uitgeverij Vantilt.

Acree, M. C. (1978). Theories of statistical inference in psychological research: A historico-critical study (Doctoral Thesis). Clark University, Worcester, MA, US.

Albani, A., Lombardo, G. P., & Proietto, M. (2014). Storia e indirizzi nell'Istituto di psicologia Sperimentale dell'Universitá di Roma da Sante De Sanctis a Mario Ponzo. In G. P. Lombardo (Ed.), *Storia e "crisi" della Psicologia scientifica in Italia* (pp. 89–171). Milano: LED.

Andersen, C., Bek-Thomsen, J., & Kjærgaard, P. C. (2012). The Money Trail: A New Historiography for Networks, Patronage, and Scientific Careers. *Isis*, *103*(2), 310–315. https://doi.org/10.1086/666357

Araujo, S. (2016). *Wundt and the Philosophical Foundations of Psychology: A Reappraisal*. Cham, Switzerland: Springer International Publishing.

Araujo, S. (2017). Toward a philosophical history of psychology: An alternative path for the future. *Theory & Psychology*, *27*(1), 87–107. https://doi.org/10.1177/0959354316656062

Atkinson, R. L., Atkinson, R. C., & Hilgard, E. R. (1983). *Introduction to psychology* (8th ed.). New York, NY: Harcourt Brace Jovanovich.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., & Bem, D. J. (1993). *Introduction to psychology* (11th ed.). Fort Worth, TX: Harcourt Brace Jovanovich.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Hilgard, E. R. (1990). *Introduction to psychology* (10th ed.). San Diego, CA: Harcourt Brace Jovanovich.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (1996). *Introduction to psychology* (12th ed.). Fort Worth, TX: Harcourt, Brace College Publishers.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (2000). *Introduction to psychology*. (C. D. Smith, Ed.) (13th ed.). Fort Worth, TX: Harcourt, Brace College Publishers.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., & Hilgard, E. R. (1987). *Introduction to psychology* (9th ed.). San Diego, CA: Harcourt Brace Jovanovich.

Bakan, D. (1966). Comments on D. Fiske's 'On the Coordination of Personality Concepts and their Measurement.' *Human Development*, *9*(1–2), 84–88. https://doi.org/10.1159/000270372

Baker, M. (2016). Statisticians issue warning over misuse of P values. *Nature News*, *531*(7593), 151. https://doi.org/10.1038/nature.2016.19503

Bakker, M., Dijk, A. van, & Wicherts, J. M. (2012). The Rules of the Game Called Psychological Science. *Perspectives on Psychological Science*, *7*(6), 543–554. https://doi.org/10.1177/1745691612459060

Barclay, A. M. (1973). Death and Rebirth in Psychology. *Contemporary Psychology*, *18*(7), 333–334. https://doi.org/10.1037/0012122

Bazerman, C. (1987). Codifying the Social Scientific Style: The APA Publication Manual as Behaviorist Rhetoric. In J. S. Nelson, A. Megill, & D. N. McCloskey (Eds.), *The Rhetoric of the Human Sciences: Language and Argument in Scholarship and Public Affairs* (pp. 257–277). Madison, WI: University of Wisconsin Press.

Bazerman, C. (1988*). Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison, WI: University of Wisconsin Press.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*(3), 407–425. https://doi.org/10.1037/a0021524

Bench, S. W., Rivera, G. N., Schlegel, R. J., Hicks, J. A., & Lench, H. C. (2017). Does expertise matter in replication? An examination of the reproducibility project: Psychology. *Journal of Experimental Social Psychology*, *68*, 181–184. https://doi.org/10.1016/j.jesp.2016.07.003

Benjafield, J. G. (2012). The long past and short history of the vocabulary of anglophone psychology. *History of Psychology*, *15*(1), 50–71. https://doi.org/10.1037/a0023386

Benjafield, J. G. (2013). The vocabulary of anglophone psychology in the context of other subjects. *History of Psychology*, *16*(1), 36–56. https://doi.org/10.1037/a0030532

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Benjamin, L. T., & VandenBos, G. R. (2006). The Window on Psychology's Literature: A History of Psychological Abstracts. *American Psychologist*, *61*(9), 941–954. http://dx.doi.org/10.1037/0003-066X.61.9.941

Bensaude-Vincent, B. (2006). Textbooks on the Map of Science Studies. *Science & Education*, *15*(7–8), 667–670. https://doi.org/10.1007/s11191-005-1243-1

Bishop, M. A., & Trout, J. D. (2004). *Epistemology and the Psychology of Human Judgment*. Oxford, UK: Oxford University Press.

Blum, A. (2017). The Literature Review as Imagined Past. *Isis*, *108*(4), 827–829. https://doi.org/10.1086/695604

Bohannon, J. (2016). Who's downloading pirated papers? Everyone. *Science*, *352*(6285), 508–512. https://doi.org/10.1126/science.352.6285.508

Borg, I., & Groenen, P. J. F. (2005). *Modern Multidimensional Scaling: Theory and Applications*. New York, NY: Springer.

Boring, E. G. (1929). *A History of Experimental Psychology* (1st ed.). New York, NY: Appleton-Century-Crofts.

Boring, E. G. (1950). *A History of Experimental Psychology* (2nd ed.). New York, NY: Appleton-Century-Crofts.

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The Two Disciplines of Scientific Psychology, or: The Disunity of Psychology as a Working Hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Eds.), *Dynamic Process Methodology in the Social and Developmental Sciences* (pp. 67–97). New York: Springer. https://doi.org/10.1007/978-0-387-95922-1_4

Bouterse, J. (2016, February 25). *Nature and history: towards a hermeneutic philosophy of historiography of science* (Doctoral Thesis). Leiden University. Retrieved from https://openaccess.leidenuniv.nl/handle/1887/38041

Bovan, K. (2016, December). Rekonstrukcija i eksperimentalna provjera koncepta točnog glasovanja [Reconstruction and an experimental verification of the concept of correct voting] (Doctoral Thesis). University of Zagreb, Zagreb.

Bridgman, P. W. (1927). *The logic of modern physics*. New York, NY: Macmillan.

Bridgman, P. W. (1938). Operational Analysis. *Philosophy of Science*, 5(2), 114–131.

Brock, A. C. (1995). Why I am not a Historian of Science. *History and Philosophy of Psychology Bulletin*, *7*(2), 27–31.

Brock, A. C. (2006). Introduction. In A. C. Brock (Ed.), *Internationalizing the History of Psychology* (pp. 1–15). New York, US: NYU Press.

Brock, A. C. (2014). What is a polycentric history of psychology? *Estudos e Pesquisas Em Psicologia*, *14*(2), 646–659.

Brock, A. C. (2017a). The new history of psychology II: Some (different) answers to Watrin's four questions. *History of Psychology*, *20*(2), 238–250. https://doi.org/10.1037/hop0000060

Brock, A. C. (2017b). The new history of psychology: Some (different) answers to Lovett's five questions. *History of Psychology*, *20*(2), 195–217. https://doi.org/10.1037/hop0000036

Brown, R. M., & Brown, N. L. (1982). Bias in Psychology and Introductory Psychology Textbooks. *Psychological Reports*, *51*(3), 1195–1204. https://doi.org/10.2466/pr0.1982.51.3f.1195

Brush, S. G. (1974). Should the History of Science Be Rated X?: The way scientists behave (according to historians) might not be a good model for students. *Science*, *183*(4130), 1164–1172. https://doi.org/10.1126/science.183.4130.1164

Burgard, D. E. (2001). Journals of the Century in Psychology. *The Serials Librarian*, *39*(3), 41–56. https://doi.org/10.1300/J123v39n03_06

Burman, J. T. (2018). Through the Looking-Glass: PsycINFO as an Historical Archive of Trends in Psychology. *History of Psychology*. https://doi.org/10.1037/hop0000082

Burman, J. T., Green, C. D., & Shanker, S. (2015). On the Meanings of Self-Regulation: Digital Humanities in Service of Conceptual Clarity. *Child Development*, *86*(5), 1507–1521. https://doi.org/10.1111/cdev.12395

Capshew, J. H. (1999). *Psychologists on the March: Science, Practice, and Professional Identity in America, 1929-1969*. Cambridge, UK: Cambridge University Press.

Carson, J. (2007). *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750-1940.* Princeton, NJ: Princeton University Press.

Chambers, C. (2017). *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*. Princeton, NJ: Princeton University Press.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist*, *49*(12), 997–1003. https://doi.org/10.1037/0003-066X.49.12.997

Cohen-Cole, J. (2007). Instituting the science of mind: intellectual economies and disciplinary exchange at Harvard's Center for Cognitive Studies. *The British Journal for the History of Science*, *40*(04), 567–597. https://doi.org/10.1017/S0007087407000283

Cohen-Cole, J. (2014). *The Open Mind: Cold war politics and the sciences of human nature*. Chicago, IL: University of Chicago Press.

Collins, H. (1992). *Changing Order: Replication and Induction in Scientific Practice*. Chicago, IL: University of Chicago Press.

Coyne, J. C. (2016). Replication initiatives will not salvage the trustworthiness of psychology. *BMC Psychology*, *4*(28), 1–11. https://doi.org/10.1186/s40359-016-0134-3

Crandall, C. S., & Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, *66*, 93–99. https://doi.org/10.1016/j.jesp.2015.10.002

Crane, D. (1972). *Invisible Colleges; Diffusion of Knowledge in Scientific Communities*. Chicago, IL: University of Chicago Press.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671–684.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*(2), 116–127.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302.

Daniel, R. S. (1992). Teaching of psychology, the journal. In A. E. Puente, J. R. Matthews, & C. L. Brewer (Eds.), *Teaching of psychology in America: A history* (pp. 433–452). Washington, D.C.: American Psychological Association.

Danziger, K. (1985). The Methodological Imperative in Psychology. *Philosophy of the Social Sciences/Philosophie Des Sciences Sociales*, *15*(1), 1–13.

Danziger, K. (1990). *Constructing the Subject: Historical Origins of Psychological Research*. Cambridge, UK: Cambridge University Press.

Danziger, K. (1994). Does the History of Psychology Have a Future? *Theory & Psychology*, *4*(4), 467–484. https://doi.org/10.1177/0959354394044001

Danziger, K. (1996). The practice of psychological discourse. In C. F. Graumann & K. J. Gergen (Eds.), *Historical dimensions of psychological discourse* (pp. 17–35). New York: Cambridge University Press.

Danziger, K. (1997). *Naming the Mind: How Psychology Found Its Language*. London, UK: SAGE Publications.

Danziger, K. (2006). Universalism and Indigenization in the History of Modern Psychology. In A. C. Brock (Ed.), *Internationalizing the History of Psychology* (pp. 208–225). New York, US: NYU Press.

Danziger, K., & Dzinas, K. (1997). How psychology got its variables. *Canadian Psychology-Psychologie Canadienne*, *38*(1), 43–48. https://doi.org/10.1037/0708-5591.38.1.43

Daston, L. (2009). Science Studies and the History of Science. *Critical Inquiry*, *35*(4), 798–813. https://doi.org/10.1086/599584

Daston, L. (2015). Simon and the Sirens: A Commentary. *Isis*, *106*(3), 669–676. https://doi.org/10.1086/683531

Daston, L., & Galison, P. (2010). *Objectivity*. New York, NY: Zone Books.

De Solla Price, D. (1963/1986). *Little Science, Big Science…and Beyond*. New York, NY: Columbia University Press.

Dehue, T. (1995). *Changing the rules: Psychology in the Netherlands, 1900-1985*. Cambridge, UK: Cambridge University Press.

Derksen, M. (submitted). Putting Popper to Work. *Theory & Psychology*.

Dominus, S. (2017, October 18). When the Revolution Came for Amy Cuddy. *The New York Times*. Retrieved from https://www.nytimes.com/2017/10/18/magazine/when-the-revolution-came-for-amy-cuddy.html

Drucker, J. (2011). Humanities Approaches to Graphical Display. *Digital Humanities Quarterly*, *5*(1). Retrieved from http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html

Eagly, A. H., Eaton, A., Rose, S. M., Riger, S., & Mchugh, M. C. (2012). Feminism and Psychology: Analysis of a Half-century of Research on Women and Gender. *American Psychologist*, *67*(3), 211–230. https://doi.org/10.1037/a0027260

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00621

Erickson, P., Klein, J. L., Daston, L., Lemov, R., Sturm, T., & Gordin, M. D. (2013). *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality*. Chicago, IL: University of Chicago Press.

Etz, A., & Vandekerckhove, J. (2016). A Bayesian Perspective on the Reproducibility Project: Psychology. *PLOS ONE*, *11*(2), e0149794. https://doi.org/10.1371/journal.pone.0149794

Fabry, G., & Fischer, M. R. (2015). Replication--The ugly duckling of science? *GMS Zeitschrift Fur Medizinische Ausbildung*, *32*(5), Doc57. https://doi.org/10.3205/zma000999

Fanelli, D. (2010a). Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *PLOS ONE*, *5*(4), e10271. https://doi.org/10.1371/journal.pone.0010271

Fanelli, D. (2010b). "Positive" Results Increase Down the Hierarchy of the Sciences. *PLOS ONE*, *5*(4), e10068. https://doi.org/10.1371/journal.pone.0010068

Fanelli, D. (2018). Opinion: Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences*, *115*(11), 2628–2631. https://doi.org/10.1073/pnas.1708272114

Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences*, *41*(2), 131–149. https://doi.org/10.1002/jhbs.20079

Feest, U. (2018). Thinking About Concepts in (Conceptual) Replication. Presented at the Reflections on Replication: Psychology's current crisis, Utrecht, Netherlands.

Ferguson, C. J., Brown, J. M., & Torres, A. V. (2016). Education or Indoctrination? The Accuracy of Introductory Psychology Textbooks in Covering Controversial Topics and Urban Legends About Psychology. *Current Psychology*, 1–9. https://doi.org/10.1007/s12144-016-9539-7

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059

Fiedler, K., & Prager, J. (2018). The Regression Trap and Other Pitfalls of Replication Science—Illustrated by the Report of the Open Science Collaboration. *Basic and Applied Social Psychology*, 0(0), 1–10. https://doi.org/10.1080/01973533.2017.1421953

Fleck, L. (1979). *The genesis and development of a scientific fact*. (F. Bradley & T. J. Trenn, Trans.). Chicago, IL: University of Chicago Press.

Flis, I. (2013, December). Interpretiraju li studenti psihologije testove značajnosti nul hipoteza dovoljno kritično? [Do psychology students interpret null hypothesis significance testing critically] (Master thesis). University of Zagreb, Zagreb, Croatia.

Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, *19*(6), 975–991. https://doi.org/10.3758/s13423-012-0322-y

Fuchs, A. H. (2000). Contributions Of American Mental Philosophers To Psychology In The United States. *History of Psychology*, *3*(1), 3–19. http://dx.doi.org/10.1037/1093-4510.3.1.3

Fuller, S. (1988). *Social Epistemology*. Bloomington, IN: Indiana University Press.

Furumoto, L. (1989). The new history of psychology. In I. S. Cohen (Ed.), *The G. Stanley Hall lecture series, Vol. 9* (pp. 9–34). Washington, DC: American Psychological Association.

Galison, P. (1987). *How Experiments End*. Chicago, IL: University of Chicago Press.

Garfield, E., Pudovkin, A. I., & Istomin, V. S. (2003). Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology*, *54*(5), 400–412. https://doi.org/10.1002/asi.10226

Gelman, A. (2016, September 21). What has happened down here is the winds have changed [Blog]. Retrieved October 10, 2016, from http://andrewgelman.com/2016/09/21/what-has-happened-down-here-is-the-winds-have-changed/

Gergen, K. J. (1973). Social Psychology as History. *Journal of Personality and Social Psychology*, *26*(2), 309–320.

Gieryn, T. F. (1999). *Cultural Boundaries of Science: Credibility on the Line*. Chicago, IL: University of Chicago Press.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*(2), 254–267. https://doi.org/10.1037/0033-295X.98.2.254

Gigerenzer, G. (1998a). Surrogates for Theories. *Theory & Psychology*, *8*(2), 195–204. https://doi.org/10.1177/0959354398082006

Gigerenzer, G. (1998b). We need statistical thinking, not statistical rituals. *Behavioral and Brain Sciences*, *21*(02), 199–200. https://doi.org/10.1017/S0140525X98281167

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, *33*(5), 587–606. https://doi.org/10.1016/j.socec.2004.09.033

Gigerenzer, G., Krauss, S., & Vitouch, O. (2004). The null ritual: What you always wanted to know about significance testing but were afraid to ask. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 391–408). Thousand Oaks: Sage.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1990). *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge, UK: Cambridge University Press.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science." *Science*, *351*(6277), 1037–1037. https://doi.org/10.1126/science.aad7243

Gillen, B. (1973). Readability and human interest scores of thirty-four current introductory psychology texts. *American Psychologist*, *28*(11), 1010–1011.

Gliner, J. A., Leech, N. L., & Morgan, G. A. (2002). Problems With Null Hypothesis Significance Testing (NHST): What Do the Textbooks Say? *The Journal of Experimental Education*, *71*(1), 83–92. https://doi.org/10.1080/00220970209602058

Goertzen, J. R. (2008). On the Possibility of Unification The Reality and Nature of the Crisis in Psychology. *Theory & Psychology*, *18*(6), 829–852. https://doi.org/10.1177/0959354308097260

Goldman, A. I. (1999). *Knowledge in a Social World*. Wotton-under-Edge, UK: Clarendon Press.

Gordin, M. D. (2015). *Scientific Babel: How science was done before and after global English.* Chicago, IL: University of Chicago Press.

Graiver, I. (2017, October 10). Probing the Boundary between Knowledge and Science in the History of Psychology: The Late Antique Roots of Introspection. Retrieved October 10, 2017, from http://www.shellsandpebbles.com/2017/10/10/probing-the-boundary-between-knowledge-and-science-in-the-history-of-psychology-the-late-antique-roots-of-introspection/

Green, C. D. (1992a). Is unified positivism the answer to psychology's disunity? *American Psychologist*, *47*(8), 1057–1058. https://doi.org/10.1037/0003-066X.47.8.1057

Green, C. D. (1992b). Of Immortal Mythological Beasts Operationism in Psychology. *Theory & Psychology*, *2*(3), 291–320. https://doi.org/10.1177/0959354392023003

Green, C. D. (2015). Why psychology isn't unified, and probably never will be. *Review of General Psychology*, *19*(3), 207–214. https://doi.org/10.1037/gpr0000051

Green, C. D. (2016). A digital future for the history of psychology? *History of Psychology*, *19*(3), 209–219. https://doi.org/10.1037/hop0000012

Green, C. D. (2018, February). *Is psychology's credibility on the verge of collapse?* Keynote address presented at the Reflections on Replication: Psychology's current crisis, Utrecht, Netherlands.

Green, C. D., & Feinerer, I. (2015). The Evolution of The American Journal of Psychology 1, 1887–1903: A Network Investigation. *The American Journal of Psychology*, *128*(3), 387–401. http://dx.doi.org/10.5406/amerjpsyc.128.3.0387

Green, C. D., Feinerer, I., & Burman, J. T. (2013). Beyond the schools of psychology 1: A digital analysis of Psychological Review, 1894–1903. *Journal of the History of the Behavioral Sciences*, *49*(2), 167–189. https://doi.org/10.1002/jhbs.21665

Green, C. D., Feinerer, I., & Burman, J. T. (2014). Beyond the schools of psychology 2: A digital analysis of Psychological Review, 1904–1923. *Journal of the History of the Behavioral Sciences*, *50*(3), 249–279. https://doi.org/10.1002/jhbs.21665

Green, C. D., Feinerer, I., & Burman, J. T. (2015a). Searching for the structure of early American psychology: Networking Psychological Review, 1894–1908. *History of Psychology*, *18*(1), 15–31. https://doi.org/10.1037/a0038406

Green, C. D., Feinerer, I., & Burman, J. T. (2015b). Searching for the structure of early American psychology: Networking Psychological Review, 1909–1923. *History of Psychology*, *18*(2), 196–204. https://doi.org/10.1037/a0039013

Griggs, R. A. (2014a). Coverage of the Stanford Prison Experiment in Introductory Psychology Textbooks. *Teaching of Psychology*, *41*(3), 195–203. https://doi.org/10.1177/0098628314537968

Griggs, R. A. (2014b). The Continuing Saga of Little Albert in Introductory Psychology Textbooks. *Teaching of Psychology*, *41*(4), 309–317. https://doi.org/10.1177/0098628314549702

Griggs, R. A. (2014c). Topical Coverage in Introductory Textbooks From the 1980s Through the 2000s. *Teaching of Psychology*, *41*(1), 5–10. https://doi.org/10.1177/0098628313514171

Griggs, R. A. (2015a). The Disappearance of Independence in Textbook Coverage of Asch's Social Pressure Experiments. *Teaching of Psychology*, *42*(2), 137–142. https://doi.org/10.1177/0098628315569939

Griggs, R. A. (2015b). The Kitty Genovese Story in Introductory Psychology Textbooks: Fifty Years Later. *Teaching of Psychology*, *42*(2), 149–152. https://doi.org/10.1177/0098628315573138

Griggs, R. A., Jackson, S. L., Christopher, A. N., & Marek, P. (1999). Introductory Psychology Textbooks: An Objective Analysis and Update. *Teaching of Psychology*, *26*(3), 182–189. https://doi.org/10.1207/S15328023TOP260304

Gross, A. G. (1990). *The Rhetoric of Science*. Cambridge, MA: Harvard University Press.

Guldi, J., & Armitage, D. (2014). *The History Manifesto*. Cambridge, UK: Cambridge University Press. Retrieved from http://historymanifesto.cambridge.org/

Hacking, I. (1965/2016). *Logic of Statistical Inference*. New York, NY: Cambridge University Press.

Hacking, I. (1996). The looping effects of human kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal Cognition: A Multidisciplinary Debate* (pp. 351–394). Oxford, UK: Oxford University Press.

Hacking, I. (2002). *Historical Ontology*. Cambridge, MA: Harvard University Press.

Hacking, I. (2006, August 17). Making Up People. *London Review of Books*, pp. 23–26.

Hale, N. (2000). *The Rise and Crisis of Psychoanalysis in the United States: Freud and the Americans, 1917-1985*. New York, NY: Oxford University Press.

Harari, H., & Jacobson, A. (1984). Teaching Psychology in the 1980s: A Content Analysis of Leading Introductory Psychology Textbooks. *Teaching of Psychology*, *11*(4), 236–237. https://doi.org/10.1177/009862838401100414

Harris, B. (1979). Whatever happened to little Albert? *American Psychologist*, *34*(2), 151–160.

Harris, B. (2011). Letting go of little Albert: Disciplinary memory, history, and the uses of myth. *Journal of the History of the Behavioral Sciences*, *47*(1), 1–17. https://doi.org/10.1002/jhbs.20470

Hartgerink, C. H. J., Wicherts, J. M., & Van Assen, M. A. L. M. (2017). Too Good to be False: Nonsignificant Results Revisited. *Collabra: Psychology*, *3*(1), 1–18. https://doi.org/10.1525/collabra.71

Hempel, C. G. (1968). Maximal Specificity and Lawlikeness in Probabilistic Explanation. *Philosophy of Science*, *35*(2), 116–133. https://doi.org/10.1086/288197

Hempel, C. G., & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, *15*(2), 135–175. https://doi.org/10.1086/286983

Henriques, G. (2003). The tree of knowledge system and the theoretical unification of psychology. *Review of General Psychology*, *7*(2), 150–182. https://doi.org/10.1037/1089-2680.7.2.150

Henriques, G. R. (2008). Special Section: The Problem of Psychology and the Integration of Human Knowledge Contrasting Wilson's Consilience with the Tree of Knowledge System. *Theory & Psychology*, *18*(6), 731–755. https://doi.org/10.1177/0959354308097255

Hilgard, E. R. (1953). *Introduction to psychology* (1st ed.). New York, NY: Harcourt, Brace.

Hilgard, E. R. (1957). *Introduction to psychology* (2nd ed.). New York, NY: Harcourt Brace and Co.

Hilgard, E. R. (1962). *Introduction to psychology* (3rd ed.). New York, NY: Harcourt, Brace & World.

Hilgard, E. R., & Atkinson, R. C. (1967). *Introduction to psychology* (4th ed.). New York, NY: Harcourt, Brace & World.

Hilgard, E. R., Atkinson, R. C., & Atkinson, R. L. (1971). *Introduction to psychology* (5th ed.). New York, NY: Harcourt Brace Jovanovich.

Hilgard, E. R., Atkinson, R. C., & Atkinson, R. L. (1975). *Introduction to psychology.* (6th ed.). New York, NY: Harcourt Brace Jovanovich.

Hilgard, E. R., Atkinson, R. L., & Atkinson, R. C. (1979). *Introduction to psychology* (7th ed.). New York. NY: Harcourt Brace Jovanovich.

Hilgard, E. R., Leary, D. E., & McGuire, G. R. (1991). The History of Psychology: A Survey and Critical Assessment. *Annual Review of Psychology*, *42*(1), 79–107. https://doi.org/10.1146/annurev.ps.42.020191.000455

Himmelstein, D. S., Romero, A. R., Levernier, J. G., Munro, T. A., McLaughlin, S. R., Tzovaras, B. G., & Greene, C. S. (2018). Research: Sci-Hub provides access to nearly all scholarly literature. *ELife*, *7*, e32822. https://doi.org/10.7554/eLife.32822

Hornstein, G. A. (1988). Quantifying Psychological Phenomena: Debates, Dilemmas, and Implications. In J. G. Morawski (Ed.), *The Rise of Experimentation in American Psychology* (pp. 1–34). New Haven, CT: Yale University Press.

Hoyningen-Huene, P. (2008). Systematicity: The Nature of Science. *Philosophia*, *36*(2), 167–180. https://doi.org/10.1007/s11406-007-9100-x

Huistra, P., & Jacobs, N. (In press). Report of "Funding Bodies in Late Modern Science" (Utrecht University, 30 November - 1 December, 2017). *Journal of the History of the Behavioral Sciences*.

Inbar, Y. (2016). Association between contextual dependence and replicability in psychology may be spurious. *Proceedings of the National Academy of Sciences*, *113*(34), E4933–E4934. https://doi.org/10.1073/pnas.1608676113

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLOS Medicine*, *2*(8), 696–701. https://doi.org/10.1371/journal.pmed.0020124

Ioannidis, J. P. A., Fanelli, D., Dunne, D. D., & Goodman, S. N. (2015). Meta-research: Evaluation and Improvement of Research Methods and Practices. *PLOS Biology*, *13*(10), 1–7. https://doi.org/10.1371/journal.pbio.1002264

Isaac, J. (2012). *Working Knowledge*. Cambridge, MA: Harvard University Press.

Jacobs, N. (2016). Summary of The History Manifesto. *Isis*, *107*(2), 311–314. https://doi.org/10.1086/687177

Jahoda, M. (1977). *Freud and the dilemmas of psychology*. New York, NY: Basic Books.

Jansz, J., & Drunen, P. V. (Eds.). (2003). *A Social History of Psychology*. Hoboken, NJ: Wiley-Blackwell.

Jevremov, T., Pajić, D., & Šipka, P. (2007). Structure of personality psychology based on cocitation analysis of prominent authors. *Psihologija*, *40*(2), 329–343.

Johnson, V. E., Payne, R. D., Wang, T., Asher, A., & Mandal, S. (2017). On the Reproducibility of Psychological Science. *Journal of the American Statistical Association*, *112*(517), 1–10. https://doi.org/10.1080/01621459.2016.1240079

Jokić, M., Zauder, K., & Letina, S. (2012). *The features of Croatian national and international scholarly productivity in social sciences, arts and humanities 1991-2005*. Zagreb, Croatia: Institut za društvena istraživanja u Zagrebu. Retrieved from http://bib.irb.hr/prikazi-rad?&rad=611719

Kahneman, D. (2012, September 26). A proposal to deal with questions about priming effects. Retrieved from http://www.nature.com/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf

Kihlstrom, J. F. (2002). In memoriam: Ernest Ropiequet Hilgard, 1904-2001. *The International Journal of Clinical and Experimental Hypnosis*, *50*(2), 95–103. https://doi.org/10.1080/00207140208410093

Kihlstrom, J. F. (2010). Great Books in Psychology. Retrieved April 11, 2018, from https://www.ocf.berkeley.edu/~jfkihlstrom/GreatBooks.htm

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, Š., Bernstein, M. J., … Nosek, B. A. (2014). Investigating Variation in Replicability. *Social Psychology*, *45*(3), 142–152. https://doi.org/10.1027/1864-9335/a000178

Klein, S. B. (2014). What can recent replication failures tell us about the theoretical commitments of psychology? *Theory & Psychology*, *24*(3), 326–338. https://doi.org/10.1177/0959354314529616

Koch, S. (1993). “Psychology” or “the psychological studies”? *American Psychologist*, *48*(8), 902–904. https://doi.org/10.1037/0003-066X.48.8.902

Koch, S., & Leary, D. E. (Eds.). (1992). *A Century of Psychology as Science*. Washington, D.C.: American Psychological Association.

Kožnjak, B. (2016). Kuhn Meets Maslow: The Psychology Behind Scientific Revolutions. *Journal for General Philosophy of Science*, 1–31. https://doi.org/10.1007/s10838-016-9352-x

Kuhn, T. S. (1962/2012). *The structure of scientific revolutions* (4th ed.). Chicago, IL: University of Chicago Press.

Kusch, M. (1995). *Psychologism: A Case Study in the Sociology of Philosophical Knowledge*. New York, NY: Routledge.

Lakatos, I. (1970). Falsificationism and the Methodology of Scientific Research Programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of Knowledge* (pp. 91–196). Cambridge, UK: Cambridge University Press.

Lakens, D., Adolfi, F. G., Albers, C. J., Anvari, F., Apps, M. A. J., Argamon, S. E., … Zwaan, R. A. (2018). Justify your alpha. *Nature Human Behaviour*, 1. https://doi.org/10.1038/s41562-018-0311-x

Lamont, P. (2012). The Making of Extraordinary Psychological Phenomena. *Journal of the History of the Behavioral Sciences*, *48*(1), 1–15. https://doi.org/10.1002/jhbs.21516

Larsen, K. R., Michie, S., Hekler, E. B., Gibson, B., Spruijt-Metz, D., Ahern, D., … Yi, J. (2017). Behavior change interventions: the potential of ontologies for advancing science and practice. *Journal of Behavioral Medicine*, *40*(1), 6–22. https://doi.org/10.1007/s10865-016-9768-0

Latour, B. (1999a). Give me a laboratory and I will raise the world. In M. Biagioli (Ed.), *The science studies reader* (pp. 258–275). London, UK: Routledge.

Latour, B. (1999b). *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York, NY: Oxford University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.

Leahey, T. H. (2002). History without the past. In W. E. Pickren & D. A. Dewsbury (Eds.), *Evolving perspectives on the history of psychology* (pp. 15–20). Washington, D.C.: American Psychological Association.

LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017). Falsifiability is not optional. *Journal of Personality and Social Psychology, 113*(2), 254–261. https://doi.org/10.1037/pspi0000106

LeBel, E. P., & Peters, K. R. (2011). Fearing the Future of Empirical Psychology: Bem's (2011) Evidence of Psi as a Case Study of Deficiencies in Modal Research Practice. *Review of General Psychology*, *15*(4), 371–379. https://doi.org/10.1037/a0025172

Letina, S. (2016). Network and actor attribute effects on the performance of researchers in two fields of social science in a small peripheral community. *Journal of Informetrics*, *10*(2), 571–595. https://doi.org/10.1016/j.joi.2016.03.007

Levelt, W. J. M., Drenth, P., & Noort, E. (2012, November 28). Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. Retrieved from https://www.tilburguniversity.edu/nl/over/profiel/kwaliteit-voorop/commissie-levelt/

Lovett, B. J. (2006). The new history of psychology: A review and critique. *History of Psychology*, *9*(1), 17–37. https://doi.org/10.1037/1093-4510.9.1.17

Lovett, B. J. (2017). For balance in the historiography of psychology: Reply to Brock (2017). *History of Psychology*, *20*(2), 218–224. https://doi.org/10.1037/hop0000053

MacMartin, C., & Winston, A. S. (2000). The rhetoric of experimental social psychology, 1930–1960: From caution to enthusiasm. *Journal of the History of the Behavioral Sciences*, *36*(4), 349–364. https://doi.org/10.1002/1520-6696(200023)36:4<349::AID-JHBS4>3.0.CO;2-X

Manicas, P. T. (2006). *A Realist Philosophy of Social Science: Explanation and Understanding*. Cambridge, UK: Cambridge University Press.

Marchel, C., & Owens, S. (2007). Qualitative research in psychology: Could William James get a job? *History of Psychology*, *10*(4), 301–324. http://dx.doi.org/10.1037/1093-4510.10.4.301

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396.

Maslow, A. H. (1954). *Motivation and personality*. New York, NY: Harper.

Maslow, A. H. (1966). *The psychology of science: A reconnaissance*. Harper & Row.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. https://doi.org/10.1037/a0039400

Mayes, R., & Horwitz, A. V. (2005). DSM-III and the revolution in the classification of mental illness. *Journal of the History of the Behavioral Sciences*, *41*(3), 249–267. https://doi.org/10.1002/jhbs.20103

McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Meehl, P. E. (1990a). Appraising and Amending Theories: The Strategy of Lakatosian Defense and Two Principles that Warrant It. *Psychological Inquiry*, *1*(2), 108–141. https://doi.org/10.1207/s15327965pli0102_1

Meehl, P. E. (1990b). Why Summaries of Research on Psychological Theories are Often Uninterpretable. *Psychological Reports*, *66*(1), 195–244. https://doi.org/10.2466/pro.1990.66.1.195

Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Cambridge, MA: Harvard University Press.

Michell, J. (1986). Measurement Scales and Statistics: A Clash of Paradigms. *Psychological Bulletin*, *100*(3), 398–407.

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, *88*(3), 355–383. https://doi.org/10.1111/j.2044-8295.1997.tb02641.x

Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge, UK: Cambridge University Press.

Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American Psychologist*, *24*(12), 1063–1075.

Mischel, W. (2008). The Toothbrush Problem. *APS Observer, 21*(11). Retrieved from https://www.psychologicalscience.org/observer/the-toothbrush-problem

Morawski, J. G. (Ed.). (1988). *The Rise of Experimentation in American Psychology*. New Haven, CT: Yale University Press.

Morawski, J. G. (1992). There Is More to Our History of Giving: The Place of Introductory Textbooks in American Psychology. *American Psychologist*, *47*(2), 161–169. http://dx.doi.org/10.1037/0003-066X.47.2.161

Morawski, J. G. (1996). Principles of selves: The rhetoric of introductory textbooks in American psychology. In C. F. Graumann & K. J. Gergen (Eds.), *Historical dimensions of psychological discourse* (pp. 145–162). Cambridge, UK: Cambridge University Press.

Morawski, J. G. (2005). Reflexivity and the psychologist. *History of the Human Sciences*, *18*(4), 77–105. https://doi.org/10.1177/0952695105058472

Morawski, J. G. (unpublished manuscript). *Replication as a Psychological Problem*. Unpublished Manuscript.

Morawski, J. G., & Agronick, G. (1991). A Restive Legacy: The History of Feminist Work in Experimental and Cognitive Psychology. *Psychology of Women Quarterly*, *15*(4), 567–579. https://doi.org/10.1111/j.1471-6402.1991.tb00431.x

Morgan, C. T. (1956). *Introduction to psychology*. New York, NY: McGraw-Hill.

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., … Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing Psychology through a Distributed Collaborative Network. *PsyArXiv*. https://doi.org/10.17605/OSF.IO/785QU

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Sert, N. P. du, … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*, 0021. https://doi.org/10.1038/s41562-016-0021

Nelson, J. S., Megill, A., & McCloskey, D. N. (1987). *The Rhetoric of the Human Sciences: Language and Argument in Scholarship and Public Affairs*. Madison, WI: University of Wisconsin Press.

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, *69*(6), 066133. https://doi.org/10.1103/PhysRevE.69.066133

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113. https://doi.org/10.1103/PhysRevE.69.026113

Newstead, S. E., & Makinen, S. (1997). Psychology Teaching in Europe. *European Psychologist*, *2*(1), 3–10. https://doi.org/10.1027/1016-9040.2.1.3

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*(2), 241–301. https://doi.org/10.1037/1082-989X.5.2.241

Noone, C. (2016, December). Mindfulness and critical thinking: Structural relations, short-term state effects, and long-term intervention effects (Doctoral Thesis). National University of Ireland, Galway, Galway.

Nosek, B. A., & Bar-Anan, Y. (2012). Scientific Utopia: I. Opening Scientific Communication. *Psychological Inquiry*, *23*(3), 217–243. https://doi.org/10.1080/1047840X.2012.692215

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, *7*(6), 615–631. https://doi.org/10.1177/1745691612459058

Nwoye, A. (2015). What is African Psychology the psychology of? *Theory & Psychology*, *25*(1), 96–116. https://doi.org/10.1177/0959354314565116

O’Brien, N. P. (2001). Journals of the Century in Education. *The Serials Librarian*, *39*(3), 95–102. https://doi.org/10.1300/J123v39n03_10

Olesko, K. M. (2006). Science Pedagogy as a Category of Historical Analysis: Past, Present, and Future. *Science & Education*, *15*(7–8), 863–880. https://doi.org/10.1007/s11191-005-2014-8

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Pashler, H., & Wagenmakers, E.-J. (2012). Editors’ Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science*, *7*(6), 528–530. https://doi.org/10.1177/1745691612465253

Pettit, M., Serykh, D., & Green, C. D. (2015). Multispecies Networks: Visualizing the Psychological Research of the Committee for Research in Problems of Sex. *Isis*, *106*(1), 121–149. https://doi.org/10.1086/681039

Pettit, M., & Young, J. L. (2017). Psychology and its publics. *History of the Human Sciences*, *30*(4), 3–10. https://doi.org/10.1177/0952695117722693

Pickren, W. E. (2009). Indigenization and the history of psychology. *Psychological Studies*, *54*(2), 87–95. https://doi.org/10.1007/s12646-009-0012-7

Pickren, W. E., & Schneider, S. F. (2005). *Psychology and the National Institute of Mental Health: A Historical Analysis of Science, Practice, and Policy*. Washington, D.C.: American Psychological Association.

Platt, J. R. (1964). Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, *146*(3642), 347–353. https://doi.org/10.1126/science.146.3642.347

Pooley, J. D. (2016). A “Not Particularly Felicitous” Phrase: A History of the “Behavioral Sciences” Label. *Serendipities: Journal for the Sociology and History of the Social Sciences*, *1*, 38–81.

Pooley, J. D., & Solovey, M. (2010). Marginal to the Revolution: The Curious Relationship between Economics and the Behavioral Sciences Movement in Mid-Twentieth-Century America. *History of Political Economy*, *42*(Suppl 1), 199–233. https://doi.org/10.1215/00182702-2009-077

Porter, T. M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.

Puente, A. E., Matthews, J. R., & Brewer, C. L. (1992). Introduction. In A. E. Puente, J. R. Matthews, & C. L. Brewer (Eds.), *Teaching of psychology in America: A history* (pp. 1–8). Washington, D.C.: American Psychological Association.

Quereshi, M. Y. (1981). Analytic Procedures for Selecting a General Psychology Textbook. *Teaching of Psychology*, *8*(3), 143–147. https://doi.org/10.1207/s15328023top0803_3

Quereshi, M. Y., & Buchkoski, J. E. (1979). Logical versus Empirical Estimates of Readability and Human Interest of General Psychology Textbooks. *Teaching of Psychology*, *6*(4), 202–205. https://doi.org/10.1207/s15328023top0604_3

Quereshi, M. Y., & Sackett, P. R. (1977). An Updated Content Analysis of Introductory Psychology Textbooks. *Teaching of Psychology*, *4*(1), 25–30. https://doi.org/10.1207/s15328023top0401_6

Quereshi, M. Y., & Zulli, M. R. (1975). A Content Analysis of Introductory Psychology Textbooks. *Teaching of Psychology*, *2*(2), 60–65. https://doi.org/10.1207/s15328023top0202_3

Quine, W. V. O. (1951). Two Dogmas of Empiricsm. *The Philosophical Review*, *60*(1), 20–43.

Quine, W. V. O. (1969). Epistemology Naturalized. In *Ontological Relativity and Other Essays* (Vol. 1, pp. 69–90). New York, NY: Columbia University Press.

Quine, W. V. O. (1991). Two Dogmas in Retrospect. *Canadian Journal of Philosophy*, *21*(3), 265–274. https://doi.org/10.1080/00455091.1991.10717246

Ratele, K. (2017). Four (African) psychologies. *Theory & Psychology*, *27*(3), 313–327. https://doi.org/10.1177/0959354316684215

Reisch, G. A. (2005). *How the Cold War Transformed Philosophy of Science: To the Icy Slopes of Logic*. Cambridge, UK: Cambridge University Press.

Retraction Watch, A. (2011, September 7). Dutch university investigating psych researcher Stapel for data fraud. Retrieved March 1, 2018, from https://retractionwatch.com/2011/09/07/dutch-university-investigating-psych-researcher-stapel-for-data-fraud/

Richards, G. (1987). Of What is History of Psychology a History? *The British Journal for the History of Science*, *20*(2), 201–211. https://doi.org/10.1017/S0007087400023748

Richards, G. (2002). The Psychology of Psychology: A Historically Grounded Sketch. *Theory & Psychology*, *12*(1), 7–36. https://doi.org/10.1177/0959354302121002

Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the Future: Three Unsuccessful Attempts to Replicate Bem's 'Retroactive Facilitation of Recall' Effect. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0033423

Rose, N. S. (1990). *Governing the Soul: The Shaping of the Private Self.* London, UK: Routledge.

Rose, N. S. (1998). *Inventing Our Selves: Psychology, Power, and Personhood*. Cambridge, UK: Cambridge University Press.

Rose, N. S. (2006). Power and subjectivity: Critical history and psychology. In Carl F. Graumann & K. J. Gergen (Eds.), *Historical Dimensions of Psychological Discourse* (pp. 103–124). Cambridge, UK: Cambridge University Press.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641.

Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review*, *18*(4), 682–689. https://doi.org/10.3758/s13423-011-0088-7

Rozeboom, W. W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, *57*(5), 416–428. https://doi.org/10.1037/h0042040

Rysiew, P. (2017). Naturalism in Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/spr2017/entries/epistemology-naturalized/

Salmon, W. C. (1999). Scientific Explanation. In M. H. Salmon, J. Earman, C. Glymour, J. G. Lennox, P. Machamer, J. E. McGuire, ... K. F. Schaffner (Eds.), *Introduction to the Philosophy of Science*. Indianapolis, IN: Hackett Publishing.

Samuels, R., Stich, S., & Bishop, M. (2002). Ending the rationality wars: How to make disputes about human rationality disappear. In R. Elio (Ed.), *Common Sense, Reasoning and Rationality* (pp. 236–268). Oxford, UK: Oxford University Press.

Schruijer, S. G. L. (2008, September). *Is the EAESP a Cold War baby? An investigation into the political context of its formation.* Paper presentation presented at the Divided Dreamworlds - The Cultural Cold War East and West, Utrecht, Netherlands.

Schruijer, S. G. L. (2012). Whatever happened to the 'European' in European social psychology? A study of the ambitions in founding the European Association of Experimental Social Psychology. *History of the Human Sciences*, *25*(3), 88–107. https://doi.org/10.1177/0952695111427365

Schweizer, G., & Furley, P. (2016). Reproducible research in sport and exercise psychology: The role of sample sizes. *Psychology of Sport and Exercise*, *23*, 114–122. https://doi.org/10.1016/j.psychsport.2015.11.005

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, *69*(1), 487–510. https://doi.org/10.1146/annurev-psych-122216-011845

Sigal, M. J., & Pettit, M. (2012). Information overload, professionalization, and the origins of the publication manual of the American Psychological Association. *Review of General Psychology, 16*(4), 357–363. https://doi.org/10.1037/a0028531

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: a key to the file-drawer. *Journal of Experimental Psychology. General*, *143*(2), 534–547. https://doi.org/10.1037/a0033242

Singal, J. (2016, October 12). Inside Psychology's 'Methodological Terrorism' Debate. *New York Magazine: Science of Us*. Retrieved from http://nymag.com/scienceofus/2016/10/inside-psychologys-methodological-terrorism-debate.html

Skinner, B. F. (1987). Laurence D. Smith. Behaviorism and logical positivism: A reassessment of the alliance. Stanford, Calif.: Stanford University Press, 1986. 416 pp. $42.50 (cloth). *Journal of the History of the Behavioral Sciences*, *23*(3), 206–210. https://doi.org/10.1002/1520-6696(198707)23:3<206::AID-JHBS2300230303>3.0.CO;2-V

Slaney, K. (2017). *Validating Psychological Constructs*. London, UK: Macmillan.

Smith, L. D. (1986). *Behaviorism and Logical Positivism: A Reassessment of the Alliance*. Stanford, CA: Stanford University Press.

Smith, R. (2007). *Being Human: Historical Knowledge and the Creation of Human Nature*. New York, NY: Columbia University Press.

Smith, R. (2013). *Between Mind and Nature: A History of Psychology*. Chicago, IL: University of Chicago Press.

Smyth, M. M. (2001a). Certainty and Uncertainty Sciences: Marking the Boundaries of Psychology in Introductory Textbooks. *Social Studies of Science, 31*(3), 389–416. https://doi.org/10.1177/030631201031003003

Smyth, M. M. (2001b). Fact-Making in Psychology: The Voice of the Introductory Textbook. *Theory & Psychology*, *11*(5), 609–636. https://doi.org/10.1177/0959354301115002

Smyth, M. M. (2004). Exploring Psychology's Low Epistemological Profile in Psychology Textbooks: Are Stress and Stress Disorders Made within Disciplinary Boundaries? *Theory & Psychology*, *14*(4), 527–553. https://doi.org/10.1177/0959354304044923

Snow, C. P. (1959). Two Cultures. *Science*, *130*(3373), 419–419. https://doi.org/10.1126/science.130.3373.419

Solovey, M. (2013). *Shaky Foundations: The Politics-Patronage-Social Science Nexus in Cold War America*. New Brunswick, NJ: Rutgers University Press.

Sommer, A. (2012). Psychical research and the origins of American psychology: Hugo Münsterberg, William James and Eusapia Palladino. *History of the Human Sciences*, *25*(2), 23–44. https://doi.org/10.1177/0952695112439376

Sommer, A. (2014). Psychical research in the history and philosophy of science. An introduction and review. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *48*, 38–45. https://doi.org/10.1016/j.shpsc.2014.08.004

Sommerey, C. M. (2015). The Ghost in the Classroom: Ernst Haeckel's Rhetoric of Evolution and Its Reverberations in German Biology Textbooks (1925-1958) (Doctoral Thesis). Universitaire Press Maastricht, Maastricht.

Spielberger, C. D. (1992). Foreword. In A. E. Puente, J. R. Matthews, & C. L. Brewer (Eds.), *Teaching of psychology in America: A history* (pp. xvii–xviii). Washington, D.C.: American Psychological Association.

Staats, A. W. (1983). *Psychology's Crisis of Disunity: Philosophy and Method for a Unified Science*. Westport, CT: Praeger.

Staats, A. W. (1991). Unified positivism and unification psychology: Fad or new field? *American Psychologist*, *46*(9), 899–912. https://doi.org/10.1037/0003-066X.46.9.899

Stam, H. J. (1992). The Demise of Logical Positivism: Implications of the Duhem-Quine Thesis for Psychology. In C. W. Tolman (Ed.), *Positivism in Psychology* (pp. 17–24). New York, NY: Springer.

Stam, H. J. (2000). Theorizing Health and Illness: Functionalism, Subjectivity and Reflexivity. *Journal of Health Psychology*, *5*(3), 273–283. https://doi.org/10.1177/135910530000500309

Stam, H. J. (2004). Unifying psychology: epistemological act or disciplinary maneuver? *Journal of Clinical Psychology*, *60*(12), 1259–1262. https://doi.org/10.1002/jclp.20069

Stam, H. J. (2013). An Integrated Transnatioanal History of Psychology - Finally! *PsycCRITIQUES*, *58*(46), 1–4.

Stam, H. J., Lubek, I., & Radtke, H. L. (1998). Repopulating social psychology texts: Disembodied "subjects" and embodied subjectivity. In B. M. Bayer & J. Shotter (Eds.), *Reconstructing the psychological subject: Bodies, practices and technologies* (pp. 153–186). London, UK: SAGE Publications.

Steinmetz, G. (2005). The Epistemological Unconsciouss of U.S. Sociology and the Transition to Post-Fordism: The Case of Historical Sociology. In J. Adams, E. S. Clemens, & A. S. Orloff (Eds.), *Remaking Modernity: Politics, History, and Sociology* (pp. 109–157). Durham, NC: Duke University Press.

Sternberg, R. J. (2005). *Unity in Psychology: Possibility Or Pipedream?* Washington, D.C.: American Psychological Association.

Steuer, F. B., & Ham, K. W. (2008). Psychology Textbooks: Examining Their Accuracy. *Teaching of Psychology*, *35*(3), 160–168. https://doi.org/10.1080/00986280802189197

Stevens, S. S. (1935a). The Operational Basis of Psychology. *The American Journal of Psychology*, *47*(2), 323–330. https://doi.org/10.2307/1415841

Stevens, S. S. (1935b). The operational definition of concepts. *Psychological Review*, *42*(6), 517–527.

Stroebe, W., & Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. *Perspectives on Psychological Science*, *9*(1), 59–71. https://doi.org/10.1177/1745691613514450

Sturm, T. (2012). The "Rationality Wars" in Psychology: Where They Are and Where They Could Go. *Inquiry*, *55*(1), 66–81. https://doi.org/10.1080/0020174X.2012.643628

Sturm, T., & Mülberger, A. (2012). Crisis discussions in psychology—New historical and philosophical perspectives. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(2), 425–433. https://doi.org/10.1016/j.shpsc.2011.11.001

Szucs, D., & Ioannidis, J. P. A. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLOS Biology*, *15*(3), e2000797. https://doi.org/10.1371/journal.pbio.2000797

Tennant, J. P., Dugan, J. M., Graziotin, D., Jacques, D. C., Waldner, F., Mietchen, D., … Colomb, J. (2017). A multi-disciplinary perspective on emergent and future innovations in peer review. *F1000Research*, *6*, 1151. https://doi.org/10.12688/f1000research.12037.1

Teo, T. (2006). *The Critique of Psychology: From Kant to Postcolonial Theory*. New York, NY: Springer Science & Business Media.

Teo, T. (2007). Local institutionalization, discontinuity, and German textbooks of psychology, 1816–1854. *Journal of the History of the Behavioral Sciences*, *43*(2), 135–157. https://doi.org/10.1002/jhbs.20220

Teo, T. (2015). Critical psychology: A geography of intellectual engagement and resistance. *American Psychologist*, *70*(3), 243–254. https://doi.org/10.1037/a0038727

Teo, T. (2017). From Psychological Science to the Psychological Humanities: Building a General Theory of Subjectivity. *Review of General Psychology*, *21*(4), 281–291. https://doi.org/10.1037/gpr0000132

Theunissen, L. T. G. (2001). Does disciplinary history matter? An epilogue. In R. Corbey & W. Roebroeks (Eds.), *Studying human origins: Disciplinary history and epistemology* (pp. 147–151). Amsterdam: Amsterdam University Press.

Tilburg Belief Systems Lab. (2017, November 25). Reanalyses of the Reproducibility Project: Psychology. Retrieved March 1, 2018, from https://tbslaboratory.com/2017/11/25/reanalyses-of-the-reproducibility-project-psychology/

Tolman, E. C. (1936). An Operational Analysis of "Demands." *Erkenntnis*, *6*, 383–392.

Trafimow, D., & Earp, B. D. (2016). Badly specified theories are not responsible for the replication crisis in social psychology: Comment on Klein. *Theory & Psychology*, *26*(4), 540–548. https://doi.org/10.1177/0959354316637136

Truijens, F. L. (2016). Do the numbers speak for themselves? A critical analysis of procedural objectivity in psychotherapeutic efficacy research. *Synthese*, 1–20. https://doi.org/10.1007/s11229-016-1188-8

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, *113*(23), 6454–6459. https://doi.org/10.1073/pnas.1521897113

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Van Eck, N. J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, *7*(3), 50–54.

Van Eck, N. J., & Waltman, L. (2014a). CitNetExplorer: A new software tool for analyzing and visualizing citation networks. *Journal of Informetrics*, *8*(4), 802–823. https://doi.org/10.1016/j.joi.2014.07.006

Van Eck, N. J., & Waltman, L. (2014b). Systematic retrieval of scientific literature based on citation relations: Introducing the CitNetExplorer tool. *Proceedings of the First Workshop on Bibliometric-Enhanced Information Retrieval (BIR 2014)*, 13–20.

Van Eck, N. J., & Waltman, L. (2014c). Visualizing Bibliometric Networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact* (pp. 285–320). New York: Springer. https://doi.org/10.1007/978-3-319-10377-8_13

Van Eck, N. J., Waltman, L., Dekker, R., & van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, *61*(12), 2405–2416. https://doi.org/10.1002/asi.21421

Vaughn-Blount, K., Rutherford, A., Baker, D., & Johnson, D. (2009). History's Mysteries Demystified: Becoming a Psychologist–Historian. *The American Journal of Psychology*, *122*(1), 117–129.

Vicedo, M. (2012). Playing the Game: Psychology Textbooks Speak Out about Love. *Isis*, *103*(1), 111–125. https://doi.org/10.1086/664982

Vrecko, S. (2010). Neuroscience, power and culture: an introduction. *History of the Human Sciences*, *23*(1), 1–10. https://doi.org/10.1177/0952695109354395

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432. https://doi.org/10.1037/a0022790

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. *Perspectives on Psychological Science*, *7*(6), 632–638. https://doi.org/10.1177/1745691612463078

Walsh, R. T. G., Teo, T., & Baydala, A. (2014). *A Critical History and Philosophy of Psychology: Diversity of Context, Thought, and Practice*. Cambridge, UK: Cambridge University Press.

Waltman, L., van Eck, N. J., & Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, *4*(4), 629–635. https://doi.org/10.1016/j.joi.2010.07.002

Wampold, B. E. (2001). *The Great Psychotherapy Debate: Models, Methods, and Findings.* Mahwah, NJ: L. Erlbaum Associates.

Warwick, A. (2003). *Masters of Theory: Cambridge and the Rise of Mathematical Physics*. Chicago, IL: University of Chicago Press.

Watrin, J. P. (2017). The ambiguous "new history of psychology": New questions for Brock (2017). *History of Psychology*, *20*(2), 225–237. https://doi.org/10.1037/hop0000057

Weidman, N. (2016a). Between the Counterculture and the Corporation: Abraham Maslow and Humanistic Psychology in the 1960s. In D. Kaiser & W. P. McCray (Eds.), *Groovy Science: Knowledge, Innovation, and American Counterculture* (pp. 109–141). Chicago, IL: University of Chicago Press.

Weidman, N. (2016b). Overcoming our mutual isolation: How historians and psychologists can work together. *History of Psychology*, *19*(3), 248–253. https://doi.org/10.1037/hop0000042

Weidman, N. (1999*). Constructing Scientific Psychology: Karl Lashley's Mind-Brain Debates*. Cambridge, UK: Cambridge University Press.

Weiten, W. (1988). Objective Features of Introductory Psychology Textbooks as Related to Professors' Impressions. *Teaching of Psychology*, *15*(1), 10–16. https://doi.org/10.1207/s15328023top1501_2

Weiten, W., & Wight, R. D. (1992). Portraits of a discipline: An examination of introductory psychology textbooks in America. In A. E. Puente, J. R. Matthews, & C. L. Brewer (Eds.), *Teaching psychology in America: A history* (pp. 453–502). Washington, D.C.: American Psychological Association.

Winegard, B. M., Winegard, B. M., & Deaner, R. O. (2014). Misrepresentations of Evolutionary Psychology in Sex and Gender Textbooks. *Evolutionary Psychology*, *12*(3), 147470491401200300. https://doi.org/10.1177/147470491401200301

Winston, A. S. (1988). Cause and Experiment in Introductory Psychology: An Analysis of R.S. Woodworth's Textbooks. *Teaching of Psychology*, *15*(2), 79–83. https://doi.org/10.1207/s15328023top1502_3

Winston, A. S. (1990). Robert Sessions Woodworth and the "Columbia Bible": How the Psychological Experiment Was Redefined. *The American Journal of Psychology*, *103*(3), 391–401. https://doi.org/10.2307/1423217

Winston, A. S. (2004). Controlling the Metalanguage. In A. C. Brock, J. Louw, & W. van Hoorn (Eds.), *Rediscovering the History of Psychology* (pp. 53–73). New York: Springer. https://doi.org/10.1007/0-306-48031-X_4

Winston, A. S., & Blais, D. J. (1996). What counts as an experiment?: A transdisciplinary analysis of textbooks, 1930-1970. *The American Journal of Psychology*, *109*(4), 599–616.

Winston, A. S., Butzer, B., & Ferris, M. D. (2004). Constructing difference: Heredity, intelligence, and race in textbooks, 1930-1970. In *Defining difference: Race and racism in the history of psychology* (pp. 199–229). Washington, DC, US: American Psychological Association.

Yong, E. (2012). Nobel laureate challenges psychologists to clean up their act. *Nature News*. https://doi.org/10.1038/nature.2012.11535

Young, J. L., & Green, C. D. (2013). An exploratory digital analysis of the early years of G. Stanley Hall's American Journal of Psychology and Pedagogical Seminary. *History of Psychology*, *16*(4), 249–268. http://doi.org/10.1037/a0033118

Appendices

Appendix A contains the article list of papers in the citation analysis for **Chapter 2.**

Appendix B contains additional information for the citation analysis in **Chapter 5**

Appendix A

The appendix lists all the publications found through the citation analysis of the data obtained from the articles retrieved from Web of Science. The publications are ranked by citation score – the number of citations the publication received. The first 40 entries in the table are visually represented in *Figure 2.1* in the chapter.

Publications in the Teaching of Psychology citation network

Title	**Year**	**DOI**	**Score**
portraits of a discipline: an examination of introductory psychology textbooks in america	1992	10.1037/10120-020	19
objective features of introductory psychology textbooks as related to professors impressions	1988	10.1207/s15328023top1501_2	18
introductory psychology textbooks: an objective analysis and update	1999	10.1207/S15328023TOP260304	18
brief introductory psychology textbooks - an objective analysis	1994	10.1207/s15328023top2103_1	16
introductory course content and goals	1998	10.1207/s15328023top2502_2	16
pedagogical aids in textbooks: do college students' perceptions justify their prevalence?	1999	10.1207/s15328023top2601_2	16

introductory textbooks and psychology's core concepts	2000	10.1207/s15328023top2701_1	15
the most frequently cited journal articles and authors in introductory psychology textbooks	1991	10.1207/s15328023top1801_2	13
similarity of introductory psychology textbooks: reality or illusion?	2001	10.1207/s15328023top2804_03	12
the introductory psychology textbook market - perceptions of authors and editors	1989	10.1207/s15328023top1602_3	11
students' perceptions of textbook pedagogical aids	1996	10.1207/s15328023top2302_8	11
introductory psychology textbooks: assessing levels of difficulty	1999	10.1207/s15328023top260401	11
the most frequently listed courses in the undergraduate psychology curriculum	1999	10.1207/S15328023TOP260303	11
the contents of introductory psychology textbooks - a follow-up	1993	10.1207/s15328023top2004_4	10
content-analysis of introductory psychology textbooks	1975	10.1207/s15328023top0202_3	9
updated content-analysis of introductory psychology textbooks	1977	10.1207/s15328023top0401_6	9
future of the introductory psychology textbook: a survey of college publishers	1997	10.1207/s15328023top2402_7	9
history from our textbooks - boring, langfeld, and weld introductory texts (1935-1948+)	1991	10.1207/s15328023top1801_9	8
university, community college, and high school students' evaluations of textbook pedagogical aids	1999	10.1207/s15328023top2601_3	8
social and abnormal psychology textbooks: an objective analysis	2000	10.1207/s15328023top2703_04	7

brief introductory psychology textbooks: a current analysis	2001	10.1207/s15328023top2801_09	7
psychology textbooks: examining their accuracy	2008	10.1080/00986280802189197	7
student use of introductory texts: comparative survey findings from two universities	2002	10.1207/s15328023top2904_13	6
pedagogical aids and student performance	2003	10.1207/s15328023top3002_01	6
using common core vocabulary in text selection and teaching the introductory course	2004		6
logical versus empirical estimates of readability and human interest of general psychology textbooks	1979	10.1207/s15328023top0604_3	5
whos who in american introductory psychology textbooks - a citation study	1985	10.1207/s15328023top1201_4	5
forty years of introductory psychology: an analysis of the first 10 editions of hilgard et al's textbook	1996	10.1207/s15328023top2303_1	5
are all of your students represented in their textbooks? a content analysis of coverage of diversity issues in introductory psychology textbooks	1997	10.1207/s15328023top2402_3	5
critical thinking in introductory psychology texts and supplements	1998		5
introductory psychology textbooks: an objective analysis update	2013	10.1177/0098628313487455	5
textbooks - problems of publishers and professors	1976		4
improving textbook selection	1985	10.1207/s15328023top1203_9	4
data graphs in introductory and upper level psychology textbooks: a content analysis	2000	10.1207/s15328023top2702_03	4
research methods textbooks: an objective analysis	2001		4

effective student use of computerized quizzes	2001	10.1207/s15328023top2804_10	4
a citation analysis of who's who in introductory textbooks	2002	10.1207/s15328023top2903_04	4
textbook selection: balance between the pedagogy, the publisher, and the student	2002		4
pedagogical aids: learning enhancers or dangerous detours?	2004	10.1207/s15328023top3103_1	4
analytic procedures for selecting a general psychology textbook	1981	10.1207/s15328023top0803_3	3
the misrepresentation of school-psychology in introduction to psychology textbooks	1981	10.1207/s15328023top0803_13	3
coverage of research ethics in introductory and social-psychology textbooks	1984		3
tales in a textbook - learning in the traditional and narrative modes	1989	10.1207/s15328023top1603_4	3
essential topics in introductory statistics and methodology courses	1997	10.1207/s15328023top2404_2	3
effects on content acquisition of signaling key concepts in text material	2003	10.1207/s15328023top3003_06	3
the most frequently cited books in introductory texts	2004		3
three decades of social psychology: a longitudinal analysis of baron and byrne's textbook	2004	10.1207/s15328023top3101_8	3
psychology textbook network	1977		2
selecting a general psychology textbook	1981	10.1207/s15328023top0804_6	2
do students remember pictures in psychology textbooks	1983	10.1207/s15328023top1001_6	2

teaching psychology in the 1980s - a content-analysis of leading introductory psychology textbooks	1984		2
the treatment of industrial organizational-psychology in introductory psychology textbooks	1984		2
faculty awareness of textbook prices	1988	10.1207/s15328023top1501_3	2
on actualizing person-centered theory - a critique of textbook treatments of rogers motivational constructs	1989	10.1207/s15328023top1601_11	2
coverage of parapsychology in introductory psychology textbooks	1991	10.1207/s15328023top1803_6	2
analysis of information about television in developmental-psychology textbooks	1992	10.1207/s15328023top1902_4	2
enhancing the psychology of memory by enhancing memory of psychology	1994	10.1207/s15328023top2103_12	2
teaching students about graphs	1995	10.1207/s15328023top2202_9	2
presentation of women and gilligan's ethic of care in college textbooks, 1970-1990: an examination of bias	1997	10.1207/s15328023top2403_2	2
supplementary books on critical thinking	1998		2
using first-person accounts to teach students about psychological disorders	2000	10.1207/s15328023top2701_9	2
useful analyses for selecting a cognitive psychology textbook	2001		2
the need for comparative textbooks: a review and research in developmental evaluation	2002	10.1207/s15328023top2902_03	2
in search of introductory psychology's classic core vocabulary	2002		2

use of primary source readings in psychology courses at liberal arts colleges	2005	10.1207/s15328023top3201_6	2
how accurately do introductory psychology textbooks present psychoanalytic theory?	2011	10.1177/0098628310390912	2
an analysis of learning objectives and content coverage in introductory psychology syllabi	2013	10.1177/0098628313487456	2
psych lite: great price, less filling	2013	10.1177/0098628313487415	2
family taboo in psychology textbooks	1978	10.1207/s15328023top0503_1	1
student feedback in writing textbooks	1980	10.1207/s15328023top0701_8	1
an analysis of the treatment of jensenism in introductory psychology textbooks	1980	10.1207/s15328023top0703_2	1
textbook evaluations by students	1983	10.1207/s15328023top1003_26	1
readability of introductory psychology textbooks - flesch versus student-ratings	1984	10.1207/s15328023top1102_8	1
the treatment of sociobiology in introductory psychology textbooks	1986	10.1207/s15328023top1301_3	1
subject emphasis in textbook photographs and journal reports in educational-psychology	1988	10.1207/s15328023top1503_21	1
schizophrenogenic parenting in abnormal-psychology textbooks	1989	10.1207/s15328023top1601_12	1
the representation of counseling versus clinical-psychology in introductory psychology textbooks	1991	10.1207/s15328023top1801_3	1
good reads in psychology - recommended books beyond the required textbook	1993	10.1207/s15328023top2003_17	1
great books in psychology - 3 studies in search of a consensus	1994	10.1207/s15328023top2102_5	1

heads and tales in introductory psychology	1996	10.1207/s15328023top2303_2	1
confronting heterosexism in the teaching of psychology	1996	10.1207/s15328023top2304_3	1
inaccurate representation of the electra complex in psychology textbooks	1997	10.1207/s15328023top2404_11	1
a valid demonstration of the missing fundamental illusion	1998	10.1207/s15328023top2501_8	1
coming to terms with the keyword method in introductory psychology: a ""neuromnemonic"" example	1998	10.1207/s15328023top2502_15	1
teaching the history of psychology: what's hot and what's not	1998	10.1207/s15328023top2503_12	1
computerized cognition laboratory	1999	10.1207/s15328023top2601_19	1
effects of day care and maternal employment: views from introductory psychology textbooks	1999		1
the portrayal of child sexual assault in introductory psychology textbooks	1999	10.1207/s15328023top260402	1
portrayals of wundt and titchener in introductory psychology texts: a content analysis	2000		1
incorporating published autobiographies into the abnormal psychology course	2001	10.1207/s15328023top2802_13	1
human sexuality textbooks: a critical look at the visual presentation of sexually explicit material	2001		1
coverage of industrial/organizational psychology in introductory psychology textbooks: an update	2002		1
operant conditioning concepts in introductory psychology textbooks and their companion web sites	2002	10.1207/s15328023top2904_04	1

using a core textbook for the introductory course	2002		1
general psychology course evaluations: differential survey response by expected grade	2003		1
textbook coverage of ethical considerations in research with children	2003		1
motion parallax: is it presented accurately in textbooks'?	2003		1
evaluation of a web site in cognitive science	2003	10.1207/s15328023top3003_11	1
core terms in undergraduate statistics	2005		1
promoting conceptual understanding of statistics: definitional versus computational formulas	2005		1
graphing psychology: an analysis of source material of graphs in introductory psychology textbooks	2005		1
effect of textbook study guides on student performance in introductory psychology	2005	10.1207/s15328023top3201_8	1
introductory psychology topics and student performance: where's the challenge?	2006	10.1207/s15328023top3303_2	1
an inclusive process for departmental textbook selection	2006	10.1207/s15328023top3304_2	1
introductory psychology student performance: weekly quizzes followed by a cumulative final exam	2007		1
classic articles as primary source reading in introductory psychology	2007		1
predicting textbook reading: the textbook assessment and usage scale	2011	10.1177/0098628310390913	1
what happened to the first ""r""?: students' perceptions of the role of textbooks in psychology courses	2011	10.1177/0098628311421319	1

brief introductory psychology textbooks: an objective analysis update	2013	10.1177/0098628313501280	1
topical coverage in introductory textbooks from the 1980s through the 2000s	2014	10.1177/0098628313514171	1
some notes on this introductory psychology textbook special section	1977	10.1207/s15328023top0401_3	0
textbook abuse	1977	10.1207/s15328023top0401_13	0
rationale for practice of randomly assigning subjects to groups - its treatment in textbooks in experimental-psychology and some suggestions	1977	10.1207/s15328023top0402_12	0
teaching undergraduate statistics with and without a textbook	1978		0
publishing a textbook - advice from authors and publishers	1978	10.1207/s15328023top0504_1	0
rear end analysis - uses of social-psychology textbook citation data	1979	10.1207/s15328023top0602_12	0
adjuncts to the textbook for an undergraduate clinical-psychology class	1980	10.1207/s15328023top0701_13	0
assessment of psychology instructors perceptions and use of textbook test-item manuals for measuring student-achievement	1981	10.1207/s15328023top0802_6	0
little albert from the viewpoint of abnormal-psychology textbook authors	1983	10.1207/s15328023top1004_14	0
textbook of psychology, 4th edition - hebb,do, donderi,dc	1988	10.1207/s15328023top1501_15	0
instructors manual - textbook of psychology, 4th edition - donderi,dc, henderson,as	1988	10.1207/s15328023top1501_15	0
cause and experiment in introductory psychology - an analysis of woodworth,r.s. textbooks	1988	10.1207/s15328023top1502_3	0

citations of distinguished scientists in introductory psychology textbooks	1989	10.1207/s15328023top1602_8	0
understanding and applying psychology through use of news clippings	1992	10.1207/s15328023top1903_8	0
integrating suicidology into abnormal-psychology classes - the revised facts on suicide quiz	1992	10.1207/s15328023top1903_9	0
ideas for teaching history and systems	1992	10.1207/s15328023top1903_18	0
applied sources for teachers of nonparametric statistics	1993	10.1207/s15328023top2001_12	0
preventing lost syllabi	1993	10.1207/s15328023top2002_13	0
how should textbooks summarize the status of parapsychology - a reply	1993	10.1207/s15328023top2003_10	0
summarizing parapsychology in psychology textbooks - a rejoinder	1993	10.1207/s15328023top2003_11	0
the science fair - a supplement to the lecture technique	1993	10.1207/s15328023top2004_8	0
an exercise for explicating and critiquing students implicit personality theories	1994	10.1207/s15328023top2103_13	0
theoretical and applied sources for teachers of structural equation modeling	1994	10.1207/s15328023top2103_15	0
inquiring minds really do want to know - using questioning to teach critical thinking	1995	10.1207/s15328023top2201_5	0
critiquing articles cited in the introductory textbook: a writing assignment	1995	10.1207/s15328023top2204_4	0
applied sources for teachers of structural equation modeling	1995	10.1207/s15328023top2204_8	0
a child panel to facilitate the instruction of child development	1996	10.1207/s15328023top2303_7	0

computer-assisted instruction as a supplement to lectures in an introductory psychology class	1996	10.1207/s15328023top2303_9	0
what monkeys can do	1997	10.1207/s15328023top2401_16	0
a life stress instrument for classroom use	1998	10.1207/s15328023top2501_15	0
capturing the fervor of cognitive psychology's emergence	1998	10.1207/s15328023top2502_17	0
a sweet way to teach students about the sampling distribution of the mean	1998	10.1207/s15328023top2503_6	0
we dream, you do: ""great"" grandmothers teach a lesson in women's changing roles	1998		0
an introduction to textbook publishing: what we did not learn in graduate school	1999	10.1207/s15328023top2601_8	0
searching for a common core: an examination of human sexuality textbook references	1999	10.1207/s15328023top2602_14	0
introductory psychology textbooks: an objective analysis and update	1999		0
positive health psychology: an interview with shelley taylor	2000	10.1207/s15328023top2704_09	0
correlational analysis and interpretation: graphs prevent gaffes	2001	10.1207/s15328023top2802_14	0
family, friends, and self: the real-life context of an abnormal psychology class	2001		0
difficulty and discriminability of introductory psychology test items	2001	10.1207/s15328023top2801_03	0
an alternative approach to the ill-defined problem of teaching problem solving	2001		0

an evaluation of industrial/organizational psychology teaching modules for use in introductory psychology	2002	10.1207/s15328023top2901_10	0
helping students gain insight into mental set	2003		0
diversity research in teaching of psychology: summary and recommendations	2003	10.1207/s15328023top3001_02	0
an argument for a laboratory in introductory psychology	2003		0
career pathway information in introductory psychology textbooks	2004		0
applying the just-in-time teaching approach to teaching statistics	2004		0
writing exercises for introductory psychology	2005		0
elaborations of introductory psychology terms: effects on test performance and subjective ratings	2005	10.1207/s15328023top3201_7	0
a seminar on scientific writing for students, postdoctoral trainees, and junior faculty	2006		0
depth and motion perceptions produced by motion parallax	2006		0
developing and presenting auditory demonstrations: two sound editor programs	2006	10.1207/s15328023top3303_9	0
the elusive definition of outliers in introductory statistics textbooks for behavioral sciences	2006		0
guiding questions enhance student learning from educational videos	2006	10.1207/s15328023top3301_7	0
avoiding confusion surrounding the phrase ""correlation does not imply causation"""	2006		0

deviations from apa style in textbook sample manuscripts	2006		0
teaching psychological science through writing	2007		0
searching under cups for clues about memory: an online demonstration	2007		0
evaluating the electronic textbook: is it time to dispense with the paper text.?	2008	10.1080/00986280701818532	0
models and exemplars of scholarship in the teaching of psychology	2008	10.1080/00986280802373908	0
the representation of applied psychology areas in introductory psychology textbooks	2008	10.1080/00986280802189130	0
the effect of online chapter quizzes on exam performance in an undergraduate social psychology course	2009	10.1080/00986280802528972	0
scholarship in teaching and learning: an interview with john mitterer	2009	10.1080/00986280802528923	0
the compleat teacher-scholar: an interview with stephen f. davis	2009	10.1080/00986280902959879	0
limited access: the status of disability in introductory psychology textbooks	2010	10.1080/00986280903426290	0
online discussion assignments improve students' class preparation	2010	10.1080/00986283.2010.488546	0
psychology ethics in introductory psychology textbooks	2011	10.1177/0098628311401583	0
benefits of student-generated note packets: a preliminary investigation of sq3r implementation	2011	10.1177/0098628311411786	0
a multi-modal active learning experience for teaching social categorization	2011	10.1177/0098628311411783	0

students learn equally well from digital as from paperbound texts	2011	10.1177/0098628311421330	0
a multisite study of learning in introductory psychology courses	2012	10.1177/0098628312450428	0
teaching generation me	2013	10.1177/0098628312465870	0
a pilot study of core topics in introductory social psychology and developmental psychology textbooks	2014	10.1177/0098628313514184	0
topical coverage in introductory psychology: textbooks versus lectures	2014	10.1177/0098628314530347	0
coverage of the stanford prison experiment in introductory psychology textbooks	2014	10.1177/0098628314537968	0
you can lead a horse to water: efficacy of and students' perceptions of an online textbook support site	2014	10.1177/0098628314537987	0
victor the wild boy as a teaching tool for the history of psychology	2014	10.1177/0098628314537977	0

Note: All publications are from the journal Teaching of Psychology, except the first one, which is from the volume Teaching of Psychology in America: A history

Appendix B

Figure 5.1 in Chapter 5 was produced by using CitNetExplorer, v. 1.0.0, an openly available program for citation analysis designed and maintained by Nees Jan van Eck and Ludo Waltman at CWTS at Leiden University - http://www.citnetexplorer.nl/

Search query

The visualization was based on the data retrieved from Clairvate Analytics database Web of Science, accessed through the library subscription at Utrecht University. I searched the Web of Science Core Collection on January 24th 2018, using the Cited Reference Search. The used query was "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises", searched in the field Cited Title.

The query produced 23 hits. They were inspected individually to ensure that they indeed refer to Raymond Nickerson's 1998 paper under the above title. In this way, two of the hits were disqualified, as show below.

Cited Reference Search
Find the articles that cite a person's work.
Step 2: Select cited references and click "Finish Search."
Hint: Look for cited reference variants (sometimes different pages of the same article are cited or papers are cited incorrectly).

View our Cited Reference Search tutorial.

CITED REFERENCE INDEX
References: 1 - 23 of 23

Page 1 of 1

Select Page | Select All* | Clear All | Finish Search

Select	Cited Author	Cited Work [SHOW EXPANDED TITLES]	Year	Volume	Issue	Page	Identifier	Citing Articles **	View Record
☑	Nickerson	REV GEN PSYCHOL	1998	2	2	175		17	
☑	Nickerson, R. S.	REV GEN PSYCHOL	1998	2		175		2	
☑	Nickerson	REV GEN PSYCHOL	1998	2		175		1	
☑	Nickerson, R	REV GEN PSYCHOL	1998	22	2	175		1	
☑	Nickerson, R.	REV GEN PSYCHOL	1998	2	2			2	
☑	Nickerson, Raymond S.	GEN PSYCHOL	1998	2		175		1	
☑	Nickerson, Raymond S	REV GEN PSYCHOL	1998	2		175		1	
☑	Nickerson, RS	REV GEN PSYCHOL	1998	2		45		3	
☑	Nickerson, R. S.	REV GEN PSYCHOL	1998	2		75		3	
☑	Nickerson, R. S.	REV GEN PSYCHOL	1998	2		174		1	
☑	Nickerson, R. S.	REV GEN PSYCHOL	1998	2		175		1	
☑	Nickerson, Raymond S.	REV GEN PSYCHOL	1998	2		175	10.1037/1089-2680.2.2.175	1348	
☑	Nickerson, Raymond S.	REV GEN PSYCHOL	1998	2		176		5	
☑	Nickerson, Raymond S.	REV GEN PSYCHOL	1998	2		177		4	
☑	Nickerson, Raymond S.	REV GEN PSYCHOL	1998	2		178		5	
☑	Nickerson, Raymond S.	REV GEN PSYCHOL	1998	2		179		1	
☑	Nickerson, Raymond S.	REV GEN PSYCHOL	1998	2		210		3	
☑	Nickerson, R. S.	REV GEN PSYCHOL	1998	2		737		1	
☐	Owad, T.	REV GEN PSYCHOL	2006	2		175		2	
☑	Raymond, N. S.	REV GEN PSYCHOL	1998	2		175		3	
☑	Nickerson Raymond, S	REV GEN PSYCHOL	1998	2	2	175		2	
☐	Sperber, Dan + [Show all authors]	REASONING AS A SOCIA	2012					1	
☑	[Anonymous]	REV GEN PSYCHOL		2		175		1	
Select	Cited Author	Cited Work	Year	Volume	Issue	Page	Identifier	Citing Articles **	View Record

Select Page | Select All* | Clear All | Finish Search

Restrict results by languages and document types:

All languages
English
Afrikaans
Arabic

All document types
Article
Abstract of Published Item
Art Exhibit Review

* "Select All" adds the first 500 matches to your cited reference search, not all matches.

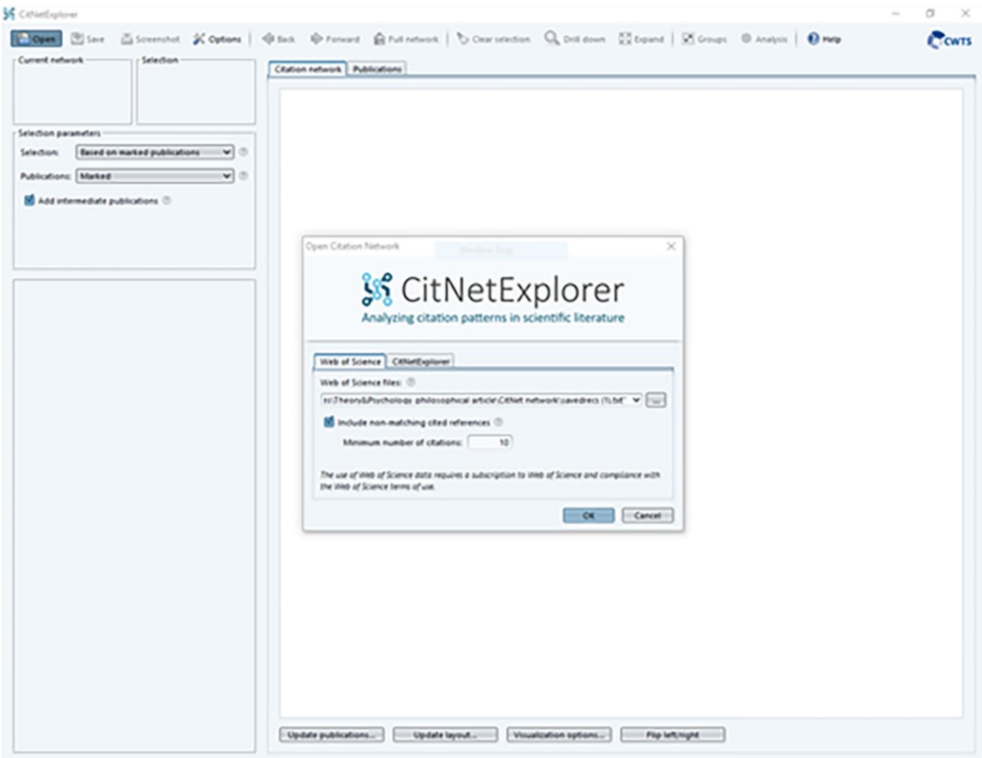
Query results

The 21 hits were cited 1202 times in WoS Core Collection. All of those citing articles were download in three files containing up to 500 full citations each. To download them, one needs to click *Save in Other File Formats -> Full Record and Cited References -> Tab-Delimited (Win).*

Producing the visualization

Step 1 – Loading WoS files into CitNetExplorer

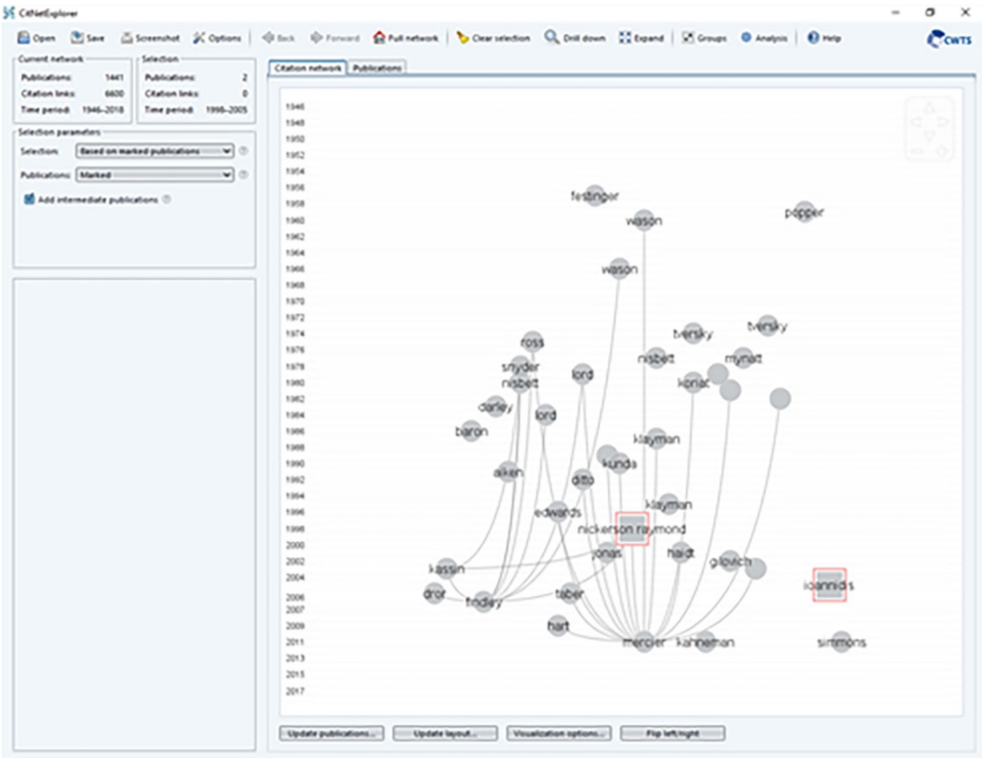
The visualization is produced by loading the files downloaded from WoS into CitNetExplorer. I will cover all the steps taken in producing the visualization so my analysis can be inspected and reproduced. Unfortunately, I cannot share the data retrieved from WoS because it is proprietary. When loading the WoS files, I chose the default minimum number of citations per paper to be included which is 10 and the option to include non-matching cited references. WoS's cited reference record is spotty at best, so it often does not include full information for each reference.



Screen for loading WoS data in Step 1

Step 2 – Drilling down to psychology's reformers

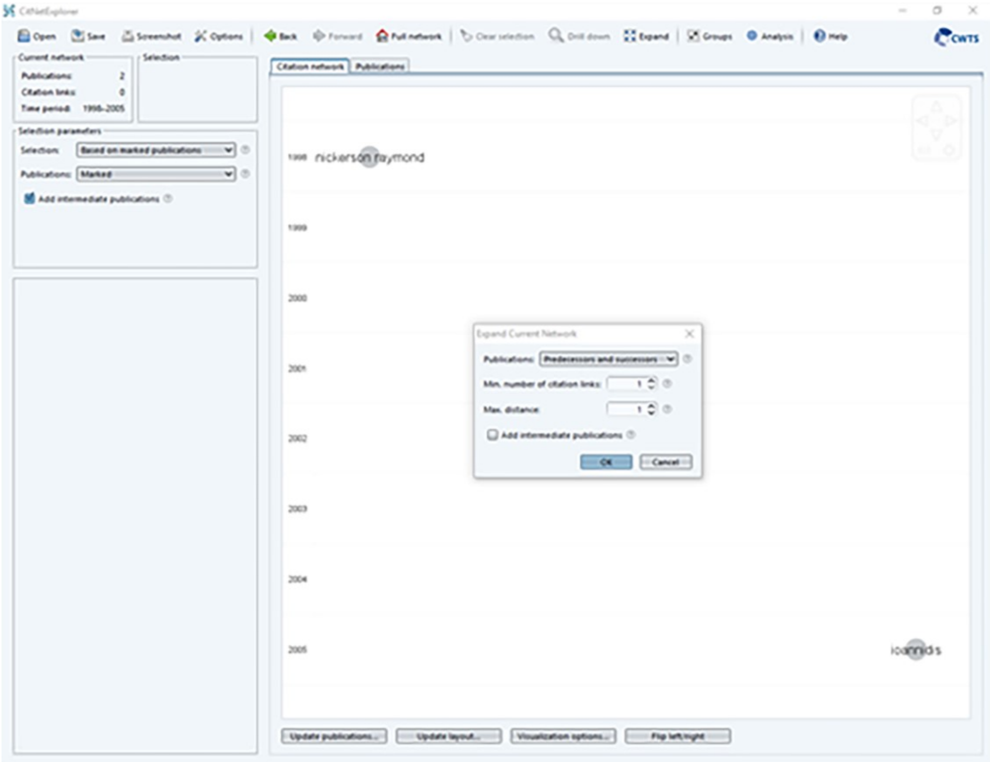
The first step produces the first citation network. The network visualizes 40 publications out of 1441 in the time period from 1946 to 2018.



The first network visualization

Multiple ways of drilling down to the parts of the citation network you are interested are possible. Keep in mind that the visualization is only a small part of the whole citation network. I will describe the steps I used to produce the visualization in the chapter. I selected two publications – Raymond Nickerson's 1998 paper *Confirmation bias: A ubiquitous phenomenon in many guises* from the *Review of General Psychology* (in the first network visualization, it is the 'nickerson raymond' red square in the middle) and the John Ioannidis 2005 paper *Why Most Published Research Findings Are False* from PLOS (the ioannidis red square on the right).

In CitNetExplorer's selection criteria, I chose to drill down 'Based on marked publications' (the above two papers). Drilling down on those two papers produced an intermediate visualization with just the two papers.

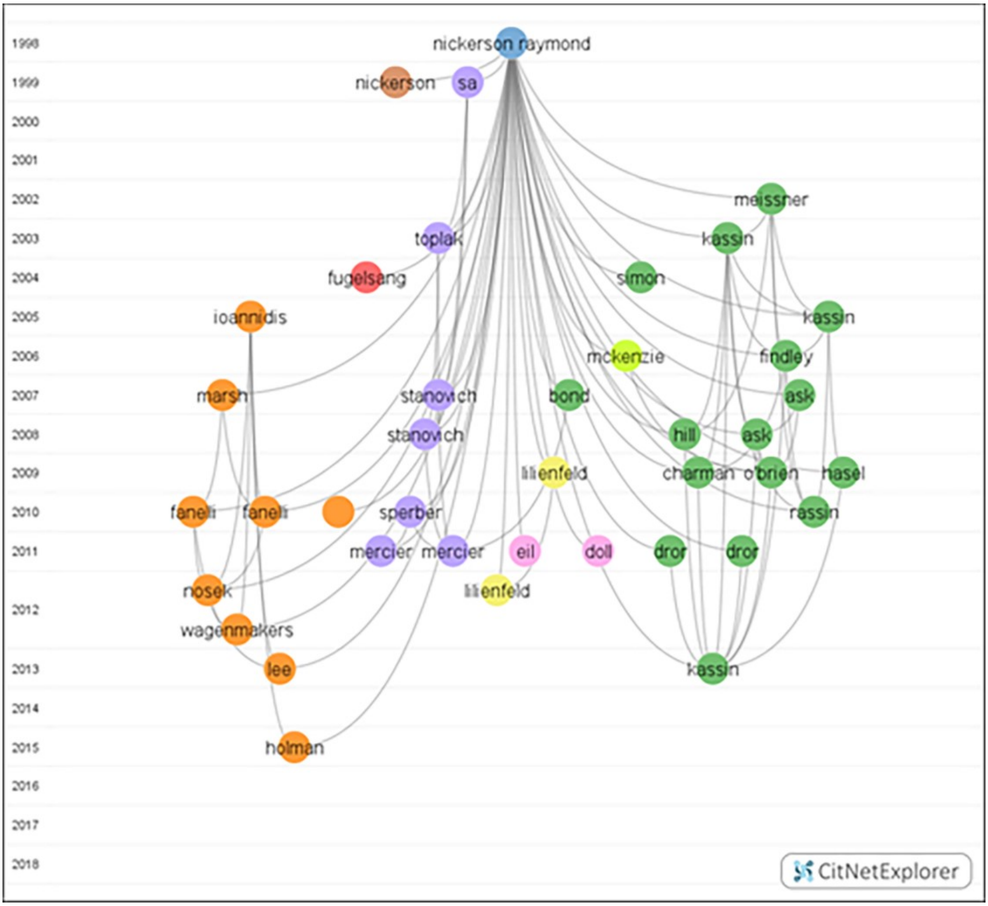


Step 2 - Expanding from the two papers

To see what kind of papers surround those two – the small group of publications relating both to Nickerson's and Ioannidis' papers – I used the CitNetExplorer Expand function. In the Expand function, I chose only to Expand into the citations that are Successors to the Nickerson and Ioannidis papers, in order to avoid the large literature that Nickerson cites in his review paper which is not interesting for my analysis. I specified the minimum number of citation links as 1 and the maximum distance as 1, to get the very closest successors to Nickerson and Ioannidis.

This Expand function produced the visualization that was used in the article. I used the default clustering of CitNetExplorer to color the groups of papers for easier inspection (click Analysis -> Clustering; the default parameters are Resolution 1.0; Minimum cluster size: 10; unselected 'Merge small clusters'; Number of random starts: 1; Number of iterations: 10; Random seed: 0).

The visualized citation network represents the 40 most cited papers in close proximity to the Nickerson and Ioannidis papers. The orange cluster is easily identified as the cluster that contains papers on science reform and the replication crisis. Considering the way we expanded from the Nickerson and Ioannidis papers, all the publications in the orange cluster necessarily cite either the Ioannidis paper or the Nickerson paper, or both (except the Ioannidis paper itself – it is actually the only paper in the visualized orange cluster that does not directly cite Nickerson. Four of the papers from the orange cluster were chosen to be discussed in detail in the article: two papers by Daniele Fanelli, one by Nosek, Spies, and Motyl, and one by Wagenmakers, Wetzels, Borsboom, van der Maas, and Kievit. They were chosen based on high citation scores.



Visualization in the chapter

The other visualized papers in the orange cluster are the following (from the top of the group to bottom, the easternmost orange bubble without a name is the Hergovich paper):

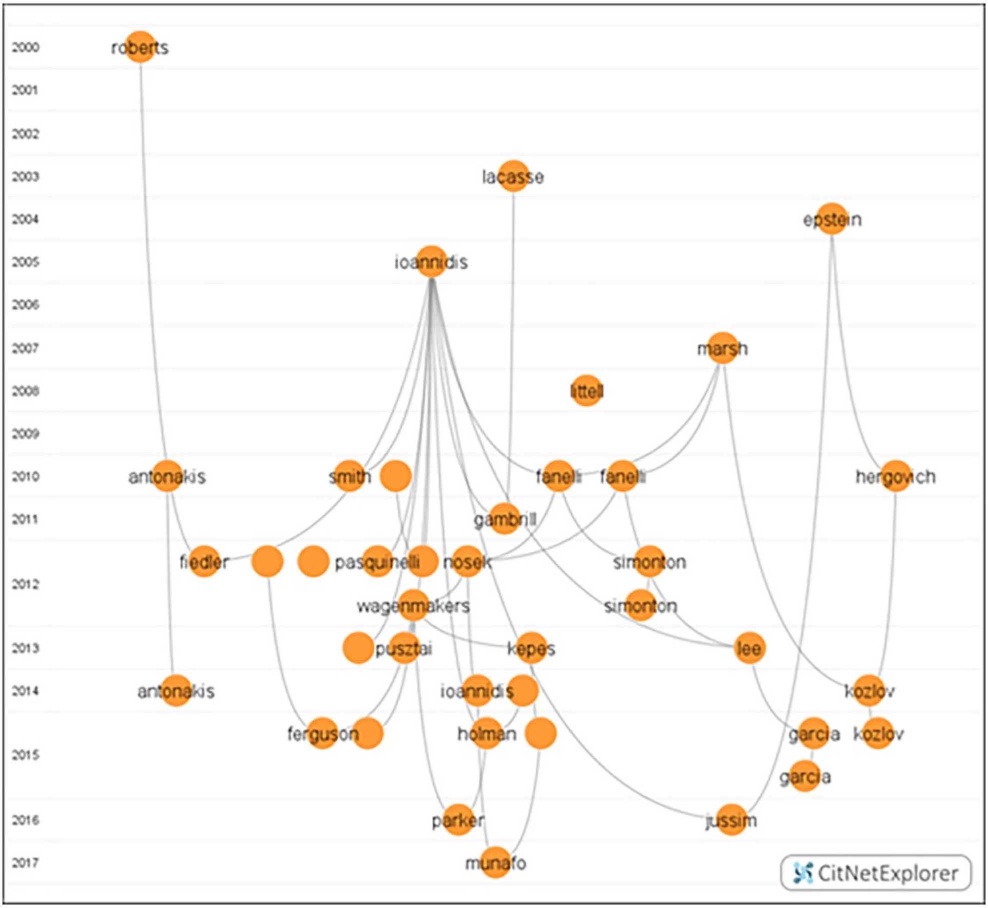
Marsh, D. M., & Hanlon, T. J. (2007). Seeing what we want to see: Confirmation bias in animal behavior research. *Ethology, 113*(11), 1089-1098.

Hergovich, A., Schott, R., & Burger, C. (2010). Biased evaluation of abstracts depending on topic and conclusion: Further evidence of a confirmation bias within scientific psychology. *Current Psychology, 29*(3), 188-209.

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the Association for Information Science and Technology, 64*(1), 2-17.

Holman, L., Head, M. L., Lanfear, R., & Jennions, M. D. (2015). Evidence of experimental bias in the life sciences: why we need blind data recording. *PLoS biology, 13*(7), e1002190.

Keep in mind that the citation network that produced the visualization is much larger and messier than what was discussed in the article. If we increase the number of visualized publications and drill down to just the orange group, we can see the full network of 83 publications in the orange cluster, for example. Inspecting this network indeed shows that the papers discussing science reform are nestled in downstream from the Nickerson's review of confirmation bias. This visualization could be further clustered, or expanded in different directions, but that goes beyond this article.



Drilled down orange cluster (all 83 publications)

The full list of publications in the orange cluster can be found in the supplementary materials on Figshare, following this link: https://doi.org/10.6084/m9.figshare.5962564.v1

Publications included in the thesis

Chapter 2 was published in the *Journal of the History of the Behavioral Sciences.*

Flis, I. (2016). Instructional Manuals of Boundary-Work: Psychology textbooks, student subjectivities, and disciplinary historiographies. *Journal of the History of the Behavioral Sciences*, 53(3), 258-278.

Chapter 4 was published in *History of Psychology*.

Flis, I. & Van Eck, N. J. (2017). Framing Psychology as a Discipline (1950-1999): A large-scale term co-occurrence analysis of scientific literature in psychology. *History of Psychology* [published online ahead of print].

Chapter 5 is currently under review in *Theory of Psychology* and a slightly different version was accepted for publication.

Flis, I. (accepted for publication). Psychologists psychologizing scientific psychology: An epistemological reading of the replication crisis. *Theory & Psychology.*

Samenvatting in het Nederlands

Het centrale onderwerp van dit proefschrift is de wetenschappelijke psychologie in de late twintigste eeuw en hoe deze zich heeft ontwikkeld. Het doel is om een globaal overzicht te bieden van een wetenschappelijke discipline die sinds de Tweede Wereldoorlog enorm is gegroeid in termen van subdisciplines, tijdschriften, onderzoekers en praktijkmensen. En toch, volgens veel psychologen zelf ontbrak het de psychologie in deze jaren aan theoretische eenheid of zelfs disciplinaire consensus over de inhoud van de psychologie. In het proefschrift wordt beargumenteerd dat de stabiliteit van de psychologie als discipline in deze periode werd bereikt door de institutionalisering van specifieke procedures voor het genereren van psychologische kennis. Aan de psychologie als wetenschap lagen geen kennisclaims, theorieën of beschrijvingen van verschijnselen ten grondslag, maar de manier waarop deze zaken geproduceerd werden. Voor psychologen kwam hun wetenschappelijke discipline tot stand door de institutionalisering van geschikte onderzoeksmethoden. Deze historische these wordt vervolgens gebruikt voor een kritische analyse van het opereren van de wetenschappelijke psychologie in het recente hervormingsdebat over de aanhoudende replicatiecrisis.

De wetenschappelijke psychologie werd de afgelopen zeventig jaar gekenmerkt door een aantal elementen die zijn geïnstitutionaliseerd door middel van tijdschriftbeleid, het universitair onderwijs en de onderzoekspraktijk. Om te beginnen is de psychologie sinds het midden van de twintigste eeuw, net als de meeste andere wetenschappelijke disciplines, enorm veramerikaniseerd. Tegen het einde van de twintigste eeuw domineerde het Engels volledig als de *lingua franca* van de wetenschap. Daarnaast hebben de Amerikaanse sociale wetenschappen in de decennia na de Tweede Wereldoorlog ongekende onderzoeksgelden ontvangen en zijn Amerikaanse wetenschappers iconen, en hun afdelingen prestigecentra geworden voor de rest van de discipline. Deze 'Amerikanisering' was niet alleen abstract of cultureel, maar manifesteerde zich ook door de institutionalisering van bepaalde praktijken in onderwijs en onderzoek.

In het proefschrift (hoofdstuk 1) identificeer ik vijf belangrijke aspecten van de wetenschappelijke psychologie die de basis vormen voor mijn verdere analyse. Deze elementen kwamen voor het eerst tot stand in de context van de Amerikaanse wetenschappelijke psychologie, ten tijde van het tanende neobehaviorisme en de cognitieve revolutie. Echter, in dezelfde periode hield deze tak van de wetenschappelijke psychologie op met exclusief Amerikaans te zijn en werd hij *het* voorbeeld voor wat het betekent om wetenschappelijk te zijn in psychologisch onderzoek, vooral wanneer dat onderzoek kwantitatief is.

Het eerste element is de filosofische opvatting 'operationalisme', die in de jaren vijftig werd voorgesteld door de natuurkundige Percy Bridgman en die werd gepopulariseerd en aangepast aan de psychologie door S.S. Stevens. Operationalisme was geen top-down filosofisch project dat de psychologie domineerde, maar een reeks van pragmatische manieren voor psychologen om hun meetpraktijken te verbinden met theoretische en/of dagelijks relevante verschijnselen. Ten tweede werd een aantal statistische procedures voor het trekken van conclusies uit data geïnstitutionaliseerd als de enige manier om kennisclaims in psychologisch onderzoek en andere kwantitatieve sociale wetenschappen te rechtvaardigen. De institutionalisering van deze inferentiële statistiek was een proces dat de psychologie in de jaren vijftig en zestig stormenderhand veroverde door tijdschriftbeleid en universitair onderwijs. Ten derde werden het operationalisme en de inferentiële statistiek samengebracht in één gemeenschappelijke taal, onder de noemer "constructvaliditeit-theorie". Constructvaliditeit-theorie werd systematisch uitgewerkt in de laatste decennia van de twintigste eeuw, een ontwikkeling die plaatsvond in de psychometrie en de vele 'toegepaste' velden in de psychologie. Deze valide constructies, die door specifieke metingen worden geoperationaliseerd, werden representatief voor *de* input die getoetst kon worden met inferentiële statistiek. Deze combinatie van operationalisme, inferentiële statistiek, en het constructie-taalgebruik werd de dominante manier waarop psychologen in de gehele discipline hun verschijnselen conceptualiseren. Ten vierde codeerden tijdschriftredacteuren en professionele organisaties de manier waarop onderzoeksrapporten worden geschreven. In deze periode werd de 'APA-stijl' een uniform format, zowel in de schrijfstijl, het soort argumenten, als de structuur van het rapport. Het had een sterk normatief effect door zijn institutionalisering in het zogeheten publicatiehandboek. Ten vijfde waren de nieuwe methodologische taal en het uniforme format zeer bevorderlijk voor het type publicatiepraktijk dat kenmerkend is voor de laatmoderne wetenschap. Wetenschappers maakten carrière door te publiceren in een steeds groter wordend aantal commerciële tijdschriften, gerangschikt op basis van prestige.

Wetenschappelijke psychologen, met hun methoden en gestandaardiseerde onderzoeksrapporten, produceerden een stortvloed aan onderzoek, waardoor de literatuurproductie van de discipline na de Tweede Wereldoorlog elk decennium in omvang verdubbelde. Velen zagen de honderdduizenden artikelen die ze publiceerden als bouwstenen van een groeiend maar nog onvoltooid corpus. Als maar genoeg studies werden toegevoegd, en als deze 'robuust' genoeg waren, d.w.z. bepaalde methodologische normen volgden, dan zou uiteindelijk een soort van zelfgeorganiseerde psychologische kennis ontstaan. Als maar genoeg puzzelstukjes geproduceerd zouden worden, dan zouden psychologen in de toekomst deze 'onvoltooide literatuur' vanzelf transformeren tot een robuuste, min of meer uniforme, theorie van de psychologie.

Hoe dominant was deze visie onder psychologen? Was het een eigenaardigheid van de Amerikaanse psychologie of van sommige subdisciplines? In het proefschrift onderzoek ik deze vragen op drie manieren. In het eerste deel (hoofdstukken 2 en 3) onderzoek ik de inleidende studieboeken die psychologen in de bestudeerde periode gebruikten om studenten te leren wat psychologie was. Studieboeken bieden een bijzondere introductie tot een discipline. Deze introductie is geschreven door de auteur van het handboek, die zich positioneert als een deskundige die de discipline uiteenzet en afbakent voor leken-studenten. Op deze manier biedt de auteur een deskundig perspectief op de discipline dat zowel algemeen is als kan rekenen op een brede consensus onder 'peers'. Ik heb in het bijzonder één handboek bestudeerd: Hilgards *Introduction to Psychology*. Ik heb gekozen voor dit populaire handboek, omdat het gebruikt is in de periode vanaf 1950 tot nu. Door het bestuderen van de introducties van de heruitgaven van Hilgards handboek toon ik aan dat de samenhang tussen de methoden van de psychologie en het proces van het 'afronden' van theorieën door middel van een 'onvoltooide literatuur' de rode draad was om de ingewikkelde verscheidenheid van psychologisch onderzoek tot een eenheid te weven. Hilgard en zijn co-auteurs beschreven de psychologie expliciet en consistent aan de hand van haar methoden. Daarnaast werd de verbinding tussen een 'onvoltooide literatuur' en een mogelijke toekomstige theorie door de auteurs zelf gelegd om het verleden en de toekomst van de discipline te duiden.

In het tweede deel (hoofdstuk 4) onderzoek ik de wetenschappelijke psychologie in dezelfde periode op het niveau van de literatuur die door de discipline gepubliceerd is. In samenwerking met Nees Jan van Eck heb ik eerder ontwikkelde scientometrische instrumenten aangepast om de Engelstalige literatuur van de psychologie die sinds 1950 gepubliceerd is op grote schaal te analyseren. We hebben de titels en abstracts van 676.393 artikelen ingeladen, die in de periode van 1950 tot 1999 geïndexeerd zijn in de APA database PsycINFO. Vervolgens hebben we automatisch de termen ontgonnen die in de titels en abstracts voorkomen en hebben we gevisualiseerd hoe deze in de tweede helft van de twintigste eeuw met elkaar samenhangen. Deze visualisaties tonen een stabiele structuur van de psychologische literatuur gedurende deze vijf decennia, waarbij de hele literatuur in tweeën was gesplitst. De ene zijde bestond uit onderzoek dat gebruikmaakte van wat psychologen gewoonlijk een "experimentele onderzoeksopzet" noemen. De andere kant bestond uit een mengelmoes van verschillende onderzoekslijnen die ze van oudsher "correlationele" psychologie noemen—een naam die Lee Cornbach, een belangrijke Amerikaanse psycholoog en psychometrist uit het midden van de twintigste eeuw, gaf aan al het onderzoek dat niet voldeed aan de strenge criteria van het Amerikaanse model voor psychologische experimenten. Zowel de "experimentele" als de "correlationele" psychologie volgden de conventies van de wetenschappelijke psychologie die ik eerder heb beschreven. We namen de stabiele structuur van de literatuur als bewijs voor de onveranderlijkheid van de eerder beschreven methodologische standaardisatie. Psychologen hebben zichzelf de taak gegeven om studies toe te voegen aan de literatuur volgens de bestaande methodologische consensus, in de hoop dat een dergelijke wetenschappelijke overproductie cumulatief zou werken en op enig moment in de toekomst zou leiden tot de totstandkoming van psychologische kennis.

Als de literatuur stabiel was en psychologen inderdaad een soort minimale consensus bereikt hebben over wat strikte wetenschappelijkheid inhoudt, waarom beweer ik dan in mijn analyse dat het resultaat van al deze vergaarde studies niet daadwerkelijk tot psychologische kennis leidt? Ik werk dit uit in het derde deel van het proefschrift (hoofdstuk 5) door de replicatiecrisis van de 21ste eeuw te bespreken. In de jaren 2010 sloten veel psychologen zich aan bij een beweging die de methodologische normen in hun discipline bekritiseerde. Zij wierpen daarbij de vraag op of de vele decennia aan onderzoek feitelijk tot een cumulatieve wetenschap hebben geleid. Het argument van de hervormers is sterk methodologisch – het gaat onder andere om het bekritiseren van de steekproeven die door onderzoekers gebruikt worden, hun slordigheid in het gebruik van statistiek, de onbedoelde en bedoelde corruptie onder uitgevers van wetenschappelijke tijdschriften, en de prikkels die het laatmoderne universitair systeem beheersen. Ik identificeer een aantal van de argumenten die door deze psychologen-hervormers opgeworpen worden en laat zien dat zij verband houden met de onveranderlijkheid van geïnstitutionaliseerde methodologische conventies die ik in mijn analyse heb beschreven. Op deze wijze bied ik (a) historische context voor de kritiek en een aantal van de claims van de hervormers, en geef ik (b) een epistemologische analyse van hun hervormingsverzoeken.

De kritische conclusie die ik trek uit mijn historische onderzoek naar de wetenschappelijke psychologie, en de filosofische kritiek die ik heb op de hervormingsbeweging in de replicatiecrisis, is als volgt: psychologen hebben, door de wijdverspreide toepassing van de wetenschappelijke psychologie, normen opgelegd aan discussies over de psychologie die geen constructieve conceptuele analyse toelaten. Door de methodologische conventies te accepteren als een raamwerk, en door slechts oppervlakkig een aantal inconsequente posities uit de reguliere wetenschapsfilosofie in te zetten als de epistemologische basis van het psychologisch onderzoek, hebben wetenschappelijke psychologen een zeer productief systeem opgezet voor het genereren van empirische claims. Echter, die kleine eilandjes van data met minimale interpretatie kunnen niet worden gesynthetiseerd tot georganiseerde kennis. Om dit te corrigeren is een historisch verfijnde filosofische analyse nodig van het operationalisme, constructvaliditeit, inferentiële statistiek en van psychologische theorie. Of we moeten besluiten tot een radicale herdefiniëring van de breed geaccepteerde doelen voor een cumulatieve wetenschap van de empirische psychologie.

Curriculum Vitae

Ivan Flis was born on October 29th, 1988 in Zagreb, Croatia. He finished the IX. general gymnasium in Zagreb (IX. gimnazija) in 2007 and the same year enrolled in a Bachelor's in psychology at the Centre for Croatian Studies of the University of Zagreb. There he finished the Bachelor's (2010) and a Master's (2013) in psychology. He wrote his Master's thesis under the supervision of dr. Iva Šverko, in which he empirically investigated the knowledge of inferential statistics among Croatian psychology students. During his undergraduate and graduate studies, he was a member of the editorial board of the Journal of European Psychology Students (JEPS), published by the European Federation of Psychology Students' Associations (EFPSA). In January 2014, he started his PhD in history and philosophy of science at the Descartes Centre for the History and Philosophy of the Sciences and the Humanities at Utrecht University, working under the supervision of prof. dr. L.T.G. Theunissen and dr. R. Abma. He finished his PhD in 2018 and is currently looking for ways to continue his research in history and philosophy of psychology.

FI Scientific Library
(formerly published as CD-β Scientific Library)

100. Flis, I. (2018). *Discipline Through Method - Recent history and philosophy of scientific psychology (1950-2018).*
99. Hoeneveld, F. (2018). *Een vinger in de Amerikaanse pap. Fundamenteel fysisch en defensie onderzoek in Nederland tijdens de vroege Koude Oorlog.*
98. Stubbé-Albers, H. (2018). *Designing learning opportunities for the hardest to reach: Game-based mathematics learning for out-of-school children in Sudan*.
97. Dijk, G. van (2018). *Het opleiden van taalbewuste docenten natuurkunde, scheikunde en techniek: Een ontwerpgericht onderzoek.*
96. Zhao, Xiaoyan (2018). *Classroom assessment in Chinese primary school mathematics education.*
95. Laan, S. van der (2017). *Een varken voor iedereen. De modernisering van de Nederlandse varkensfokkerij in de twintigste eeuw.*
94. Vis, C. (2017). *Strengthening local curricular capacity in international development cooperation.*
93. Benedictus, F. (2017). *Reichenbach: Probability & the A Priori. Has the Baby Been Thrown Out with the Bathwater?*
92. Ruiter, Peter de (2016). *Het Mijnwezen in Nederlands-Oost-Indië 1850-1950.*
91. Roersch van der Hoogte, Arjo (2015). *Colonial Agro-Industrialism. Science, industry and the state in the Dutch Golden Alkaloid Age, 1850-1950.*
90. Veldhuis, M. (2015). *Improving classroom assessment in primary mathematics education.*
89. Jupri, Al (2015). *The use of applets to improve Indonesian student performance in algebra.*
88. Wijaya, A. (2015). *Context-based mathematics tasks in Indonesia: Toward better practice and achievement.*
87. Klerk, S. (2015). *Galen reconsidered. Studying drug properties and the foundations of medicine in the Dutch Republic ca. 1550-1700.*
86. Krüger, J. (2014). *Actoren en factoren achter het wiskundecurriculum sinds 1600.*
85. Lijnse, P.L. (2014). *Omzien in verwondering. Een persoonlijke terugblik op 40 jaar werken in de natuurkundedidactiek.* Utrecht University, Utrecht.

84. Weelie, D. van (2014). *Recontextualiseren van het concept biodiversiteit. Utrecht University, Utrecht.*

83. Bakker, M. (2014). *Using mini-games for learning multiplication and division: a longitudinal effect study.*

82. Ngô Vũ Thu Hăng (2014). *Design of a social constructivism-based curriculum for primary science education in Confucian heritage culture.*

81. Sun, Lei (2014). *From rhetoric to practice: enhancing environmental literacy of pupils in China.*

80. Mazereeuw, M. (2013). *The functionality of biological knowledge in the workplace. Integrating school and workplace learning about reproduction.*

79. Dierdorp, A. (2013). *Learning correlation and regression within authentic contexts.*

78. Dolfing, R. (2013). *Teachers' Professional Development in Context-based Chemistry Education. Strategies to Support Teachers in Developing Domain-specific Expertise.*

77. Mil, M.H.W. van (2013). *Learning and teaching the molecular basis of life.*

76. Antwi, V. (2013). *Interactive teaching of mechanics in a Ghanaian university context.*

75. Smit, J. (2013). *Scaffolding language in multilingual mathematics classrooms.*

74. Stolk, M.J. (2013). *Empowering chemistry teachers for context-based education. Towards a framework for design and evaluation of a teacher professional development programme in curriculum innovations.*

73. Agung, S. (2013). *Facilitating professional development of Madrasah chemistry teachers. Analysis of its establishment in the decentralized educational system of Indonesia.*

72. Wierdsma, M. (2012). *Recontextualising cellular respiration.*

71. Peltenburg, M. (2012). *Mathematical potential of special education students.*

70. Moolenbroek, A. van (2012). *Be aware of behaviour. Learning and teaching behavioural biology in secondary education.*

69. Prins, G.T., Vos, M.A.J. & Pilot, A. (2011). *Leerlingpercepties van onderzoek & ontwerpen in het technasium.*

68. Bokhove, Chr. (2011). *Use of ICT for acquiring, practicing and assessing algebraic expertise.*

67. Boerwinkel, D.J. & Waarlo, A.J. (2011). *Genomics education for decision-making. Proceedings of the second invitational workshop on genomics education, 2-3 December 2010.*

66. Kolovou, A. (2011). *Mathematical problem solving in primary school.*

65. Meijer, M. R. (2011). *Macro-meso-micro thinking with structure-property relations for chemistry. An explorative design-based study*.

64. Kortland, J. & Klaassen, C. J. W. M. (2010). *Designing theory-based teaching-learning sequences for science. Proceedings of the symposium in honour of Piet Lijnse at the time of his retirement as professor of Physics Didactics at Utrecht University*.

63. Prins, G. T. (2010).*Teaching and learning of modelling in chemistry education. Authentic practices as contexts for learning*.

62. Boerwinkel, D. J. & Waarlo, A. J. (2010). *Rethinking science curricula in the genomics era. Proceedings of an invitational workshop*.

61. Ormel, B. J. B. (2010). *Het natuurwetenschappelijk modelleren van dynamische systemen. Naar een didactiek voor het voortgezet onderwijs*.

60. Hammann, M., Waarlo, A. J., & Boersma, K. Th. (Eds.) (2010). *The nature of research in biological education: Old and new perspectives on theoretical and methodological issues – A selection of papers presented at the VIIth Conference of European Researchers in Didactics of Biology*.

59. Van Nes, F. (2009). *Young children's spatial structuring ability and emerging number sense*.

58. Engelbarts, M. (2009*). Op weg naar een didactiek voor natuurkunde-experimenten op afstand. Ontwerp en evaluatie van een via internet uitvoerbaar experiment voor leerlingen uit het voortgezet onderwijs*.

57. Buijs, K. (2008). *Leren vermenigvuldigen met meercijferige getallen*.

56. Westra, R. H. V. (2008). *Learning and teaching ecosystem behaviour in secondary education: Systems thinking and modelling in authentic practices*.

55. Hovinga, D. (2007). *Ont-dekken en toe-dekken: Leren over de veelvormige relatie van mensen met natuur in NME-leertrajecten duurzame ontwikkeling*.

54. Westra, A. S. (2006). *A new approach to teaching and learning mechanics*.

53. Van Berkel, B. (2005). *The structure of school chemistry: A quest for conditions for escape*.

52. Westbroek, H. B. (2005). *Characteristics of meaningful chemistry education: The case of water quality*.

51. Doorman, L. M. (2005). *Modelling motion: from trace graphs to instantaneous change*.

50. Bakker, A. (2004). *Design research in statistics education: on symbolizing and computer tools*.

49. Verhoeff, R. P. (2003). *Towards systems thinking in cell biology education*.

48. Drijvers, P. (2003). *Learning algebra in a computer algebra environment. Design research on the understanding of the concept of parameter*.

47. Van den Boer, C. (2003). *Een zoektocht naar verklaringen voor achterblijvende prestaties van allochtone leerlingen in het wiskundeonderwijs*.
46. Boerwinkel, D.J. (2003). *Het vormfunctieperspectief als leerdoel van natuuronderwijs. Leren kijken door de ontwerpersbril.*
45. Keijzer, R. (2003). *Teaching formal mathematics in primary education. Fraction learning as mathematising process*.
44. Smits, Th. J. M. (2003). *Werken aan kwaliteitsverbetering van leerlingonderzoek: Een studie naar de ontwikkeling en het resultaat van een scholing voor docenten*.
43. Knippels, M. C. P. J. (2002). *Coping with the abstract and complex nature of genetics in biology education – The yo-yo learning and teaching strategy*.
42. Dressler, M. (2002). *Education in Israel on collaborative management of shared water resources.*
41. Van Amerom, B.A. (2002). *Reinvention of early algebra: Developmental research on the transition from arithmetic to algebra*.
40. Van Groenestijn, M. (2002). *A gateway to numeracy. A study of numeracy in adult basic education.*
39. Menne, J. J. M. (2001). *Met sprongen vooruit: een productief oefenprogramma voor zwakke rekenaars in het getallengebied tot 100 – een onderwijsexperiment*.
38. De Jong, O., Savelsbergh, E.R. & Alblas, A. (2001). *Teaching for scientific literacy: context, competency, and curriculum.*
37. Kortland, J. (2001). *A problem-posing approach to teaching decision making about the waste issue.*
36. Lijmbach, S., Broens, M., & Hovinga, D. (2000). *Duurzaamheid als leergebied; conceptuele analyse en educatieve uitwerking*.
35. Margadant-van Arcken, M. & Van den Berg, C. (2000). *Natuur in pluralistisch perspectief – Theoretisch kader en voorbeeldlesmateriaal voor het omgaan met een veelheid aan natuurbeelden*.
34. Janssen, F. J. J. M. (1999). *Ontwerpend leren in het biologieonderwijs. Uitgewerkt en beproefd voor immunologie in het voortgezet onderwijs*.
33. De Moor, E. W. A. (1999). *Van vormleer naar realistische meetkunde – Een historisch-didactisch onderzoek van het meetkundeonderwijs aan kinderen van vier tot veertien jaar in Nederland gedurende de negentiende en twintigste eeuw*.
32. Van den Heuvel-Panhuizen, M. & Vermeer, H. J. (1999). *Verschillen tussen meisjes en jongens bij het vak rekenen-wiskunde op de basisschool – Eindrapport MOOJ-onderzoek*.

31. Beeftink, C. (2000). *Met het oog op integratie – Een studie over integratie van leerstof uit de natuurwetenschappelijke vakken in de tweede fase van het voortgezet onderwijs*.
30. Vollebregt, M. J. (1998). *A problem posing approach to teaching an initial particle model*.
29. Klein, A. S. (1998). *Flexibilization of mental arithmeticsstrategies on a different knowledge base – The empty number line in a realistic versus gradual program design*.
28. Genseberger, R. (1997). *Interessegeoriënteerd natuur- en scheikundeonderwijs – Een studie naar onderwijsontwikkeling op de Open Schoolgemeenschap Bijlmer*.
27. Kaper, W. H. (1997). *Thermodynamica leren onderwijzen*.
26. Gravemeijer, K. (1997). *The role of context and models in the development of mathematical strategies and procedures*.
25. Acampo, J. J. C. (1997). *Teaching electrochemical cells – A study on teachers' conceptions and teaching problems in secondary education*.
24. Reygel, P. C. F. (1997). *Het thema 'reproductie' in het schoolvak biologie*.
23. Roebertsen, H. (1996). *Integratie en toepassing van biologische kennis – Ontwikkeling en onderzoek van een curriculum rond het thema 'Lichaamsprocessen en Vergift'*.
22. Lijnse, P. L. & Wubbels, T. (1996). *Over natuurkundedidactiek, curriculumontwikkeling en lerarenopleiding*.
21. Buddingh', J. (1997). *Regulatie en homeostase als onderwijsthema: een biologie-didactisch onderzoek*.
20. Van Hoeve-Brouwer G. M. (1996). *Teaching structures in chemistry – An educational structure for chemical bonding*.
19. Van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*.
18. Klaassen, C. W. J. M. (1995). *A problem-posing approach to teaching the topic of radioactivity*.
17. De Jong, O., Van Roon, P. H. & De Vos, W. (1995). *Perspectives on research in chemical education.*
16. Van Keulen, H. (1995). *Making sense – Simulation-of-research in organic chemistry education*.
15. Doorman, L. M., Drijvers, P. & Kindt, M. (1994). *De grafische rekenmachine in het wiskundeonderwijs*.
14. Gravemeijer, K. (1994). *Realistic mathematics education*.
13. Lijnse, P. L. (Ed.) (1993). *European research in science education*.
12. Zuidema, J. & Van der Gaag, L. (1993). *De volgende opgave van de computer*.

11. Gravemeijer, K, Van den Heuvel Panhuizen, M., Van Donselaar, G., Ruesink, N., Streefland, L., Vermeulen, W., Te Woerd, E., & Van der Ploeg, D. (1993). *Methoden in het reken-wiskundeonderwijs, een rijke context voor vergelijkend onderzoek*.

10. Van der Valk, A. E. (1992). *Ontwikkeling in Energieonderwijs*.

9. Streefland, L. (Ed.) (1991). *Realistic mathematics education in primary schools*.

8. Van Galen, F., Dolk, M., Feijs, E., & Jonker, V. (1991). *Interactieve video in de nascholing reken-wiskunde.*

7. Elzenga, H. E. (1991). *Kwaliteit van kwantiteit*.

6. Lijnse, P. L., Licht, P., De Vos, W. & Waarlo, A. J. (Eds.) (1990). *Relating macroscopic phenomena to microscopic particles: a central problem in secondary science education*.

5. Van Driel, J. H. (1990). *Betrokken bij evenwicht*.

4. Vogelezang, M. J. (1990). *Een onverdeelbare eenheid*.

3. Wierstra, R. F. A. (1990). *Natuurkunde-onderwijs tussen leefwereld en vakstructuur*.

2. Eijkelhof, H. M. C. (1990). *Radiation and risk in physics education*.

1. Lijnse, P. L. & De Vos, W. (Eds.) (1990). *Didactiek in perspectief.*

Universiteit Utrecht

Discipline Through Method investigates the disciplinary formation of scientific psychology in the second part of the twentieth century. In the period since the 1950s, research methods in scientific psychology were institutionalized across the varied communities of experimental, animal, educational, social, clinical, and applied psychologists. In the thesis, the epistemological implications of this institutionalization are discussed through the lens of existing historical and philosophical scholarship on scientific psychology's methods. Methods and disciplinary formation of psychology are approached in three ways: through textbooks, journals, and psychologists' debates in the wake of the 2010s replication crisis. Textbooks were investigated as instructional manuals of disciplinary boundary-work, which provide expert psychologists with a platform for their broad views on the nature of psychology as a science. Journals were approached by data-mining more than half a million titles/abstracts of journal articles retrieved from the APA's database *PsycINFO*. The journals and the textbooks indicate a methodological core to how psychological research was conducted during the whole period. That methodological core is critically discussed by focusing on the replication crisis debates and the views psychologists-reformers express on the current state and future reform of their disciplinary conventions..