BRILL

NUNCIUS 38 (2023) 553–582

NUNCIUS
brill.com/nun

From Postal Scale to Psychological Apparatus

A History of Experimental Psychology Through the Reconstruction of Peirce and Jastrow's "On Small Differences of Sensation" (1885)

Claudia Cristalli | ORCID: 0000-0001-7552-0965
Tilburg University, Tilburg, The Netherlands
c.cristalli@tilburguniversity.edu

Rebecca L. Jackson | ORCID: 0000-0003-2199-1045
Indiana University, Bloomington, IN, USA
jacksrel@iu.edu

Received 30 April 2022 | Accepted 28 June 2023 |
Published online 16 August 2023

Abstract

This paper describes our reconstruction of the apparatus used in C.S. Peirce and Joseph Jastrow's 1885 psychophysical experiment, "On Small Differences of Sensation" and how it relates to persistent questions in scientific theories of measurement. We situate Peirce and Jastrow's work in the broader context of nineteenth-century discussions about the status of psychology as a science and emphasize the role of measurement and experiment in determining that status. Through our re-enactment of the experiment, we analyze the experiment's methodology, which features blinding, randomization, noise control, and a technology for data recording. Our claims are illustrated by images of historical and reconstructed apparatuses, as well as graphs of the data collected in our reproduction. Our findings exemplify the relevance of history of psychology for contemporary philosophy of science and theories of measurement, and defend experimental reconstructions as a particularly fertile method of inquiry in the history of psychophysical sciences.

Keywords

history of psychology – experimental reconstruction – history of measurement

Published with license by Koninklijke Brill NV | DOI:10.1163/18253911-bja10066
 | ISSN: 0394-7394 (print) 1825-3911 (online)

1 Introduction

This paper deals with the reconstruction and replication of a nineteenth-century experiment in sensory discrimination, carried out by classical pragmatist Charles S. Peirce and by psychologist Joseph Jastrow. The task of rebuilding a disappeared instrument involved the re-discovery of networks of experimental psychologists and of their influence, while its operation involved an in-depth engagement with the authors' methods and practices. Our findings in this article contribute to a richer understanding of the history of psychology, philosophy of measurement, and to the relevance of instrument reconstruction for contemporary research in the history and philosophy of science.

In 1885, Charles Sanders Peirce (1839–1914), lecturer in logic at Johns Hopkins University, and Joseph Jastrow (1863–1944), a doctoral student at the same institution, published "On Small Differences of Sensation." In this paper, they claimed that some consequences of the most famous "law" of experimental psychology, Fechner's psychophysical law, did not withstand close experimental examination. In particular, Peirce and Jastrow opposed the idea of portraying the threshold of discrimination between two stimuli [Unterschiedschwelle]—i.e., the point where the difference between two stimuli becomes utterly undistinguishable—as a fixed unit to measure sensation.[1] To start, Peirce and Jastrow did not accept the universal validity of Fechner's finding: a discrimination task accomplished under uniform conditions in a dedicated setting could not easily speak for the countless discrimination tasks performed at every moment in our daily life.[2] Furthermore, Peirce and Jastrow criticised the

1 In his *Elemente der Psychophysik*, Fechner states: "Mit der Thatsache der Schwelle hängt die Folgerung zusammen, dass eine Empfindung um so weiter von der Merklichkeit entfernt bleibt, je mehr der Reiz unter seinen Schwellenwerth sinkt." [With the fact of the threshold [in perception] comes the following consequence, that a sensation remains so far removed from consciousness, the more the stimulus sinks under its threshold value]. Gustav Theodor Fechner, *Elemente Der Psychophysik* (Leipzig: Breitkopf und Härtel, 1860), 15. All translations from German are by Claudia Cristalli. Following Johann Friedrich Herbart (1776–1841), Fechner assigns a numerical value—preceded by the sign minus—to those sensations, however this remains a mathematical representation of sensations under threshold (whose value is correlated with the value of their stimulus) rather than the expression of an experimental investigation of under-threshold sensations. We thank an anonymous reviewer for pushing us to clarify this point.

2 Charles Sanders Peirce and Joseph Jastrow, "On Small Differences of Sensation," *National Academy of Sciences*, no. 5 (1885): 75–85, reprinted in Charles Sanders Peirce, *Writings of Charles S. Peirce: A Chronological Edition*, vol. 5, *1884–1886*, ed. Christian J.W. Kloesel, Nathan Houser, André De Tienne, Ursula Niklas, Aleta Houser, Cathy L. Clark, and Max H. Fisch (Bloomington and Indianapolis: Indiana University Press, 1993), 122. Hereafter, we adopt the

assumption that, simply because we are not able to tell the difference between two stimuli, we would therefore not be *feeling* anything.[3]

According to their findings, sensation occurred continuously even below Fechner's supposed threshold, meaning that we still perceive something even when we are not conscious of it. If this were the case, sensation would be better modelled by the statistical "law of errors" rather than by an empirically determined threshold (to be discussed in Section 3). Peirce and Jastrow's conclusion implied a major methodological shift for experimental psychology, in that it challenged the way in which sensation was measured. Crucially, a new experimental design and apparatus had to be introduced so that, if sensation was indeed continuous past the apparent threshold of perception, the psychophysicist could capture it as such.

In what follows, we explore the problem of measuring in experimental psychology through the re-creation and analysis of Peirce and Jastrow's experiment in sensory discrimination. This experiment is situated in the tradition of classical experimental psychology, in that it investigates the (passive) discrimination of pressure on the fingertip; furthermore, it has been noted by modern interpreters for its methodological innovations, including the introduction of blinding and randomization.[4] The Peirce and Jastrow experiment has been studied in connection with allegations of "psychologism" in Peirce's philosophy[5] or to compare Peirce's theory of perception with modern signal detection theory.[6] However, its role in the history of experimental psychology and in particular its contribution to the problem of measuring sensations (and psychical phenomena more generally) remains unexplored.[7] In charting this

standard abbreviation for Peirce's works: "W" (*Writings*) followed by volume number, page number (e.g. W5, 122).

3 Peirce and Jastrow, "On Small Differences," W5, 123.

4 Ian Hacking, "Telepathy: Origins of Randomization in Experimental Design," *Isis* 79, no. 3 (1988): 427–451.

5 Claudia Cristalli, "Experimental Psychology and the Practice of Logic." *European Journal of Pragmatism and American Philosophy* IX, no. 1 (2017): 0–24. https://doi.org/10.4000/ejpap.1006 (accessed May 18, 2023).

6 Peter J. Behrens, "Peirce's Psychophysics: Then and Now," in *Charles S. Peirce and the Philosophy of Science: Papers from the Harvard Sesquicentennial Congress*, ed. Edward C. Moore (Tuscaloosa: University of Alabama Press, 1993), 309–318.

7 Experiments of fine sensory discriminations have also been used to investigate (in scientifically accepted ways) psychical claims, such as those of telepathic communication. See e.g. James Grier Miller, "Discrimination without Awareness," *The American Journal of Psychology* 52, no. 4 (1939): 562–578, footnote 1: the work was supported by a grant from the Maria E. McMaster Fund for Psychic Research. The possible relevance of results on fine sensory discrimination for psychical research had been suggested already by Peirce and Jastrow

new territory, our aim is to contribute also to a more nuanced understanding of the 1885 experiment and to highlight the benefits of replicating past scientific experiments.[8]

The paper is articulated in four sections and a conclusion. Section 2 sets up the philosophical problem of measurement in psychology from Fechner's 1860 *Elemente der Psychophysik* [Elements of Psychophysics] to manuals of experimental psychology in the early twentieth century. This broad framework allows us to highlight the merits of Peirce and Jastrow's contribution and the complexity of the problem at hand. Section 3 introduces the historical figures of Peirce and Jastrow and their role in the history of psychology, focusing on the mystery of their missing apparatus. Interestingly, the visual evidence that furnished the blueprint for our work of reconstruction also testifies to the apparatus' wider distribution in the growing network of academic psychology laboratories. Finally, Section 4 brings our re-enactment of the 1885 experiment to the fore, together with our data analysis and its role in how we came to understand Peirce and Jastrow's collection and treatment of their data. In the Conclusion, we address the relevance of Peirce and Jastrow's apparatus for the history and philosophy of experimental psychology and for its theory of measurement.

2 Measurement: Between Physics and Psychology

What does it mean to measure something? Is measurement necessary for scientific knowledge? And is there a substantial difference between physical and psychological matters, such that only the first ones are suitable for measurement? These questions, and particularly the latter two, were crucial to the development of psychophysics and to the nineteenth century discussion over the possibility of psychology as a science. The answers to these questions are still debated today among philosophers of science and practicing scientists alike.[9] In particular, Ann-Sophie Barwich and Hasok Chang, in their recent

themselves at the end of their paper: "The insight of females as well as certain "telepathic" phenomena may be explained in this way." Peirce and Jastrow, "On Small Differences," W5, 135.

8 The literature on the reconstruction of historical experiments is vast, although only a small fraction of it is dedicated to replicating experiments in psychology. For an overview of the current debate, see Hasok Chang, "How Historical Experiments Can Improve Scientific Knowledge and Science Education: The Cases of Boiling Water and Electrochemistry," *Science and Education* 20, no. 3 (2011): 317–341. https://doi.org/10.1007/s11191-010-9301-8 (accessed May 18, 2023).

9 Some interesting contributors to this discussion include: Derek C. Briggs, *Historical and Con-*

discussion of sensory measurement in contemporary studies of olfactory discrimination, argue that the difficulties exhibited by sensory measurements are not unique to the specific task at hand, but rather characterize measurement across the board:

> Though we would not deny the difficulties of sensory measurements, we want to argue that these difficulties are not unique to sensory measurements. There is the admitted complication that human subjects can serve as both the object and the instrument in sensory measurements, but the epistemological issues that plague sensory measurements are the same as those in which philosophers of science have become increasingly interested when analyzing any kind of measurement as scientific practice.[10]

In the beginning of the nineteenth century, it was not clear at all that psychology needed measurement to become a science. Following Kant's requirements of *constructability*—either a priori, like in the case of geometry, or a posteriori, like in chemical analysis—and systematicity, measurement was not strictly required for something to be a science.[11] Thus, Johann Friedrich Herbart's a pri-

ceptual Foundations of Measurement in the Human Sciences (New York: Routledge, 2022), https://doi.org/10.3166/ds.7.537--550 (accessed May 18, 2023); Ann-Sophie Barwich and Hasok Chang, "Sensory Measurements: Coordination and Standardization," *Biological Theory* 10, no. 3 (2015): 200–211, https://doi.org/10.1007/s13752-015-0222-2 (accessed May 18, 2023); and Joel Michell, "Is Psychometrics Pathological Science?" *Measurement: Interdisciplinary Research & Perspective* 6, no. 1–2 (2008): 7–24, https://doi.org/10.1080/15366360802035489 (accessed May 18, 2023).

10 Barwich and Chang, "Sensory Measurement," 201.

11 The backdrop to most discussions on the possibility of psychology to be a science is Kant's denial that psychology can ever become one: see Immanuel Kant, *Metaphysical Foundations of Natural Science*, ed. Michael Friedman (Cambridge: Cambridge University Press, 2004), 7. On this and related claims, see also Paul Guyer, *Kant and the Claims of Knowledge* (Cambridge: Cambridge University Press, 1987); Abhaya C. Nayak and Eric Sotnak, "Kant on the Impossibility of the 'Soft Sciences'," *Philosophy and Phenomenological Research* 55, no. 1 (1995): 133–151; Thomas Sturm, *Kant Und Die Wissenschaften Vom Menschen* (Paderborn: Mentis Verlag, 2009); Katharina T. Kraus, "The Soul as the 'Guiding Idea' of Psychology: Kant on Scientific Psychology, Systematicity, and the Idea of the Soul," *Studies in History and Philosophy of Science Part A* 71 (2018): 77–88; Katherina Kraus, "Quantifying Inner Experience? Kant's Mathematical Principles in the Context of Empirical Psychology," *European Journal of Philosophy* 24, no. 2, 2013: 331–357. After Kant, Jakob Fries and Hermann von Helmholtz were instrumental for the reception of Kant's philosophy as (also) a psychology, as detailed in Frederik C. Beiser, *The Genesis of Neo-Kantianism, 1796–1880* (Oxford: Oxford University Press, 2014), 24–88. Besides Fries and Helmholtz, a notable attempt to construct psychology as a science in the Kantian sense of the *Metaphysical Foundations* is found in Johann F. Herbart's *Lehrbuch zur Psychologie* (Königsberg

ori description of the mind as a system of forces, whose interaction explains how certain sensations emerge above the threshold of consciousness while others are pulled below it, counted as a model of scientific psychology.[12] Fechner revolutionized both psychology and our broader understanding of science by introducing *measurement* into the picture.

The second son of a German pastor, Gustav Theodor Fechner (1801–1887) initially enrolled in medicine, but soon discovered that his interests lay in mathematics, physics, and studies of perception. His professor of mathematics, Karl Brandan Mollweide (1774–1825), "was equally interested in color perception" and "gained renown for his critique of Goethe's color theory [in 1810];"[13] his professor of anatomy was Ernst Heinrich Weber (1795–1878), who in his 1834 book *De Tactu* ['On Sense of Touch'] described the relationship between stimulus and sensation in terms of the increase in the stimulus necessary to provoke a just noticeable difference in sensation.[14] Thus, from the very beginning of his career Fechner was exposed to research in the field of perception, although sometimes disguised as physiology or optics.

According to Heidelberger, Fechner's attention for measurement came from his engagement with French experimental physics and chemistry, which were then becoming quantitatively treated.[15] He not only translated works that

und Leipzig: August Wilhelm Unzer, 1816), translated by Margaret K. Smith as *A Text-Book in Psychology* (New York: D. Appleton and Company, 1891). For a discussion of Herbart's position in the Neo-Kantian canon, see Beiser, *The Genesis*, 91–141, and esp. 135–136 for Herbart's metaphysical foundation of psychology. These debates—old and new—are partially accounted for by of the transcendental idea of "soul"; (2) mixing psychological consideration with formal logic; and (3) the possibility of psychology as a science "properly so called," i.e. mathematically constructible. See Thomas Sturm, "Kant on Empirical Psychology: How Not to Investigate the Human Mind," in *Kant and the Sciences*, ed. Eric Watkins (Oxford: Oxford University Press, 2001, 164–196), 163.

12 Michael Heidelberger, *Nature from Within: Gustav Theodor Fechner and His Psychophysical World View*, trans. Cynthia Klohr (Pittsburgh: University of Pittsburgh Press, 2004), 31.

13 Heidelberger, *Nature from Within*, 21.

14 See *De Tactu*, in *De Pulsu, Resorptione, Auditu et Tactu. Annotationes Anatomicae et Physiologicae* (Lipsiae: prostat apud C.F. Koehler, 1834) and *Der Tastsinn und das Gemeingefühl* (Leipzig: Verlag von Wilhelm Engelmann, [1846] 1905). Both works are translated in Ernst Heinrich Weber, *E.H. Weber On The Tactile Senses*, ed. David J. Murray and Helen E. Ross (London: Psychology Press, 1996). Weber did not formulate this relationship as an equation but only stated it discursively in his two classical essays (*On the Tactile Senses*, 126–127 and 210–212, respectively). The mathematical formulation of this relation is due to Fechner in his *Elemente der Psychophysik* [Elements of Psychophysics] (Leipzig: Druck und Verlag von Breitkopf und Härtel, 1860). See David J. Murray and Helen E. Ross, "Introduction," in Weber, *On The Tactile Senses*, 9.

15 Heidelberger emphasizes Fechner's engagement with French physics and chemistry as an

became seminal for the development of mathematical physics in Germany,[16] but also performed experiments himself characterized by "a strict quantifying method" and a great amount of data produced.[17] Heidelberger underscores how such a quantifying attitude was uncommon in the German-speaking States at the time, and emphasized the pioneering role of Fechner in proposing a quantitative treatment of physical phenomena. In a sense, the historical account so far bears testimony to Chang and Barwich's intuition that epistemological problems concerning measurement in physics and psychology are congruous.[18] For the contemporaries of Fechner, measuring presented conceptual challenges not only in the domain of sensation, but also in biology, chemistry, and physics (particularly when studying electricity). From this perspective, the great contribution of Fechner is not so much in the introduction of measurement into psychology, but rather in the broader introduction of measurement in German natural science.

Once measurement was introduced in the natural scientist's practice, Fechner could use the world of physical stimuli as a way into the (heretofore) imponderable world of psychical sensations. In 1851 Fechner wrote programmatically:

> If mathematical psychology is feasible at all (which I believe to be the case), then in my opinion we shall have to take as a basis for computation the physical phenomena to which mental phenomena are connected because they provide an immediate starting point for computation and a well-defined measure. […] we have to spell out, on the basis of our approach, a definite, mathematical dependency relation between the two

antidote to his previous beliefs in the natural philosophies of Friedrich W.J. Schelling and Lorenz Oken (Heidelberger, *Nature from Within*, 27). Nonetheless, Marco Giovanelli points out that Schelling's 1797 essay *Ideen zu einer Philosophie der Natur* [*Ideas for a Philosophy of Nature*] incorporates Kant's analysis of sensation as constituted by *degrees* as found in the "Anticipations of Perception" (*Critique of Pure Reason*, B 209). See Marco Giovanelli, *Reality and Negation—Kant's Principle of Anticipations of Perception: An Investigation of Its Impact on the Post-Kantian Debate* (Dortrecht: Springer, 2011), 83.

16 Fechner's translation of Jean-Baptiste Biot's *Précis Élémentaire de Physique Expérimentale* (Paris: Chez Deterville, Libraire, Rue Hautefeuille, 1817) appeared in the same year; see Fechner, *Lehrbuch der Experimental-Physik oder Erfahrungs-Naturlehre* (Leipzig: Verlag von Leopold Voß, 1817). Fechner also translated Louis-Jacques Thenard's *Traité de Chimie Élémentaire, Théorique et Pratique* (Paris: Chez Crochard, Libraire, Cloître St. Benoît, 1813): Fechner, *Lehrbuch der theoretischen und praktischen Chemie* (Leipzig: Verlag von Leopold Voß, 1825).

17 Heidelberger, *Nature from Within*, 28.

18 Barwich and Chang, "Sensory Measurements," 202.

> [mental and physical domains], which—despite the absence of a direct, accurate measure of phenomena in the mental realm—is open to empirical confirmation […]. … we have to extend the computation, which is based on the principle of this dependency, to the quality of mental phenomena, just as computational physics has been extended to the quality of colors and sounds, and to do this in a way that is coherent with the latter …[19]

A chief obstacle in the measurement of mental phenomena is that, it seems, even if we can always order them along a scale, we cannot necessarily measure them by comparing them to a fixed standard. However, if each mental phenomenon admits of a parallel physical one, then at least an indirect measure of mental phenomena can be gathered via a measurement of their physical parallel. With Fechner, the notion of a threshold of consciousness, which had been defined purely mathematically by Herbart in 1816, gained a new experimental value. In 1860, Fechner published his *Elemente der Psychophysik* [Elements of Psychophysics], which is today recognized as the founding text of psychophysics. In the *Elements*, the possibility of measurement became reality with the introduction of the famous "Fechner's law" linking the difference between external stimuli's intensity with differences in sensation. Psychology had become not just a "mathematical," but an experimental science. Fechner's approach of using differences between physical stimuli as proxies for differences in inner sensation was there to stay.

Much later, and in the context of North-American psychology, Edward Bradford Titchener (1867–1927) codified the use of *differences in sensation* as a unit for measuring sensation in his *Experimental Psychology: A Manual of Laboratory Practice* (1905), whose second volume was dedicated to "Quantitative Experiments." As he stated, "The only thing that we can measure is the distance between two sensations or sense points, and to do this we must have our unit step or unit distance."[20]

However, the "unit step" or unit of measure of sensation was not easy to find. Still in 1914, Guy Montrose Whipple, in his *Manual of Mental and Physical Tests*, cautioned against drawing univocal conclusions from a psychological test since "there is, at the present time, scarcely a single mental test that can be

19 Gustav Theodor Fechner, "Outline of a New Principle of Mathematical Psychology (1851)," *Psychological Research* 49, no. 4 (1987): 203. https://doi.org/10.1007/BF00309027 (accessed May 18, 2023).

20 Edward Bradford Titchener, *Experimental Psychology: A Manual of Laboratory Practice*, vol. 2 (New York: Macmillan and Co., 1905), xxv.

applied unequivocally *as a psychical measuring-rod*."[21] Although "mental testing" is generally different from the project of measuring the relation between the intensity of perceived sensation and the intensity of a physical stimulus, the issue was still that of finding the external, measurable trait that would allow a reliable albeit indirect measure of the internal psychical phenomenon.[22] As Whipple put it,

> The fact is we have not agreed upon methods of procedure; *we too often do not know what we are measuring*; and we too seldom realize the astounding complexity, variety and delicacy of form of our psychical nature.[23]

The tone of these remarks makes it seem as if 80 years of experimental practice had done little to ease psychologists' discomfort with indirect measurement. Whipple went on to describe 30 tests and many more variations, from "Anthropometric Tests" to tests of the physical, motor and sensory capacity of humans, to attention and perception tests. Throughout his manual, measurements and instruments for measuring abound in spite of cautionary words and a lingering uncertainty about their reliability.

Our recreation of Peirce and Jastrow's 1885 experiment takes an insider's road into the problems of measurement that characterized psychology in its early years and that keep being relevant for psychometric and experimental psychology today.[24] It does so by bringing into focus the role of the experimental apparatus (which facilitates and at the same time constrains measurement), the experimental protocol (which aims to guarantee the reliability of the measurements performed), and the data analysis, with its technology for representation and interpretation. From a historical perspective, Peirce and Jastrow's 1885 paper "On Small Differences of Sensation" recapitulates the crucial difficulties and results of the initial developments of psychology as a quantitative

21 Guy Montrose Whipple, *Manual of Mental and Physical Tests* (Baltimore: Warwick & York, Inc., 1914), 3; emphasis added.

22 In fact, many "mental tests" kept measuring physical reactions to stimuli (such as "endurance" and "strength") but interpreted them as physical counterparts of corresponding moral or intellectual qualities (such as "will," "persistency," "determination," and such). See John A. Popplestone and Marion White McMpherson, *An Illustrated History of American Psychology* (Akron, Ohio: Akron University Press, 1994), 68–69.

23 Whipple, *Manual of Mental and Physical Tests*, 3–4; emphasis added.

24 For recent discussions (with historical insight) into today's problems of measurement in psychology, see Denny Borsboom, "True Scores," in *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics* (Cambridge: Cambridge University Press, 2005), 11–48; Briggs, *Historical and Conceptual Foundations*.

science: the task at hand is a discrimination of pressures; the sense under investigation is the sense of touch; the law under examination is Fechner's law, and Fechner's measurement-supported theory of a threshold in sensation. Finally, the use of a dedicated instrument for the experiment represents an important consequence of Fechner's program and is aimed at overcoming some weaknesses of his early experimental design.

3 Peirce, Jastrow and a Fairbanks Scale

The 1885 paper "On Small Differences of Sensation" is the result of a collaboration between Charles S. Peirce, at the time temporary lecturer in philosophy at Johns Hopkins University, and Joseph Jastrow, who enrolled as a PhD student in 1882. In a retrospective remark, Jastrow declared that "It was Charles S. Peirce, one of the most exceptional minds that America has produced, who stimulated me most directly."[25] As we will see below, it is likely that Jastrow also played a determinant role in the design of the apparatus for their experiment.

Jastrow is today considered among the half-forgotten figures of early North American psychology.[26] The son of Polish Jewish immigrants, he arrived in the United States at three years old.[27] He would go on to found the first department of experimental psychology at the University of Madison, Wisconsin in 1888 and to work extensively on perceptual illusions, including the famous duck-rabbit Gestalt image. As for Peirce, in spite of some efforts to promote his relevance for experimental psychology, he remains mostly known by philosophers as the founder (with William James) of pragmatism.[28] Peirce often identified

25 Joseph Jastrow, "Joseph Jastrow," in *A History of Psychology in Autobiography*, vol. 1, ed. Carl Murchison, (New York: Russell & Russell, 1961), 135.

26 For a biographical sketch of Joseph Jastrow see Arthur L. Blumenthal, "The Intrepid Joseph Jastrow," in *Portraits of Pioneers in Psychology*, ed. Gregory A. Kimble, Michael Wertheimer, and Charlotte L. White (Washington, DC: American Psychological Association, 1991); Thomas C. Cadwallader, "Origins and Accomplishments of Joseph Jastrow's 1888-Founded Chair of Comparative Psychology at the University of Wisconsin," *Journal of Comparative Psychology* 101 (1987): 231–236; Alexandra Lee Levin, "The Jastrows in Madison: A Chronicle of University Life, 1888–1900," *The Wisconsin Magazine of History* 46, no. 4 (1963): 243–256; Clark L. Hull, "Joseph Jastrow: 1863–1944," *The American Journal of Psychology* 57, no. 4 (1944): 581–585. Jastrow's own autobiographical sketch is in Jastrow, "Joseph Jastrow," 135–162.

27 Hull, "Joseph Jastrow," 581.

28 See Thomas C. Cadwallader, "Charles S. Peirce (1839–1914): The First American Experimental Psychologist," *Journal of the History of the Behavioral Sciences* 10, no. 3 (1974): 291–298. In general, scholars writing on Peirce's engagement in experimental psychology tend to

himself as a logician, although, with the exception of his temporary position as a lecturer at Johns Hopkins (1879–1884),[29] he mostly worked as a "scientific specialist"[30] in astronomy, geodesy, and chemistry.[31]

As in the case of Fechner, color studies were one of Peirce's first contacts with the experimental study of sensation. They were prompted by his work at the Harvard Observatory in the years 1872–1875, where he was measuring the brightness of stars. Peirce's first published psychology paper, "Note on the Sensation of Color," appeared both in the U.S. and in the U.K.[32] The influence of Peirce's early studies on color sensation can be seen in his 1885 experiment with Jastrow, in which observations in color discrimination are used as controls for the results of their experiment on the pressure sense, and as warranty for the generalizability of such results from one sense to perceptual processes at large.[33]

According to Jastrow, "The first psychological investigation made at Johns Hopkins University was likewise undertaken at Peirce's suggestion."[34] As we articulate in detail below, this psychological experiment was already striking from a methodological point of view, in so far as it was the first experiment in which blinding and randomization were thoroughly implemented.[35] Additionally, as Jastrow proudly recalled, "*The pressure balance devised for the investigation is the forerunner of all the improved pressure-balances since employed.*"[36]

look at it from the perspective of his philosophy: the question is often about the influence of psychology on his philosophy (especially his logic), and not about his contribution to experimental psychology more broadly. See Justin Humphreys, "Subconscious Inference in Peirce's Epistemology of Perception," *Transactions of the Charles S Peirce Society* 55, no. 3 (2019): 326–346; Claudia Cristalli, "Experimental Psychology and the Practice of Logic," *European Journal of Pragmatism and American Philosophy* 9, no. 1 (2017): 0–25; Francesco Bellucci, "Logic, Psychology, and Apperception: Charles S. Peirce and Johann F. Herbart," *Journal of the History of Ideas* 76, no. 1 (2015): 69–91; Vincent Colapietro, "Notes For a Sketch of a Peircean Theory of the Unconscious," *Transactions of the Charles S. Peirce Society* 31, no. 3 (1995): 482–506.

29 See Christopher D. Green, "Johns Hopkins's First Professorship in Philosophy: A Critical Pivot Point in the History of American Psychology," *American Journal of Psychology* 120, no. 2 (2007): 303–323.

30 Peirce, "Lecture on Logic," W4, 380: "The scientific specialists—pendulum swingers and the like—are doing a great and useful work …".

31 For a short but informative *Chronology* of Peirce's life and works, see Kloesel et al. ed., "Chronology," W4, xvii–xviii.

32 Peirce, "Note on the Sensation of Color," W3, 211–216.

33 Peirce and Jastrow, "On Small Differences," W5, 134, footnote 8.

34 Jastrow, "Joseph Jastrow," 135.

35 Hacking, "Telepathy," 427–451.

36 Jastrow, "Joseph Jastrow," 136; emphasis added.

Indeed, it was so much of a forerunner that, in 1885, the term "pressure balance" does not appear. The apparatus is described as "an adaptation of a Fairbanks post-office scale":

> The apparatus used was an adaptation of a "Fairbanks" post-office scale; upon the end of the beam of which was fixed a square enlargement (about one-half inch square), with a flat top, which served to convey the pressure to the finger [...].[37]

The relevance of such instruments for the history of experimental psychology cannot be overstated. As Andrew Capshew perceptively reports, "The enduring motif in the story of modern psychology is neither a person nor even an event but a place—the experimental laboratory."[38] Officially, G. Stanley Hall established the first laboratory of experimental psychology at Johns Hopkins in 1883, but Peirce and Jastrow's experiment happened outside of it. Situated at the threshold of institutionalized psychology, Peirce and Jastrow nonetheless adhered to the ideals of scientific psychology that would justify the cult of the laboratory as the centerpiece of modern psychology's identity, such as a physics-like apparatus and methodology.[39]

As mentioned in the Introduction, the aim of Peirce and Jastrow in 1885 was to challenge Fechner's psychophysical law and replace it with the "law of errors." This did not come out of disregard for Fechner's work. In fact, Peirce had been aware of Fechner's research for more than 15 years, and he had implemented Fechner's scale in his photometric observations (conducted in the early 1870s and finally published in 1878).[40] In a letter to his father Benjamin, dated July 1868, Charles wrote:

> Here is Fechner's [Elemente der] *Psychophysik*. See vol. I pp. 72 and 93 et seq. *He says he practised the experiment of saying which of two slightly differing weights is the heavier for an hour a day for several years & that his results agreed with the method of least squares*. He promises to pub-

37 Peirce and Jastrow, "On Small Differences," W5, 128.

38 James H. Capshew, "Psychologists on Site: A Reconnaissance of the Historiography of the Laboratory," *American Psychologist* 47, no. 2 (1992): 132.

39 Tweney, "Whatever Happened to the Brass and Glass? The Rise of Statistical Instruments in Psychology, 1900–1950," in *Thick Description and Fine Texture: Studies in the History of Psychology*, ed. David Baker (Akron: The University of Akron Press, 2003). See also Ryan D. Tweney, "Studies in Historical Replication in Psychology I: Introduction," *Science and Education* 17, no. 5 (2008): 467–475.

40 Peirce, *Photometric Researches*, W3, 7.

> lish his experiments in another book which has never appeared as far as I can learn. Concerning "Schwelle" or the point where the perceptibility of a stimulus begins see pp. 238–300 ….[41]

It is precisely to investigate whether perception is better modelled with the "method of least squares" (an application of the law of errors)[42] or with the notion of a "Schwelle" [threshold] in sensation that Peirce would set up the 1885 experiment. The law of errors and the notion of a threshold in perception (also referred to as "just noticeable differences," jnd, in sensation) are two incompatible models: in one case, perception is continuous even below consciousness; in the other case, perception either occurs or not, regardless of the possible continuous change in the stimulus. Figures 1 and 2 graphically show our theoretical reconstruction of how the data would look if either the method of least squares or the notion of "Schwelle" were true of perception. In Peirce and Jastrow's words:

> If there be a least perceptible difference [i.e. a threshold of perception], then when two excitations differing by less than this [i.e. below threshold] are presented to us, and we are asked to judge which is the greater, we ought to answer wrong as often as right in the long run. Whereas, if the theory of least squares is correct, we not only ought to answer right oftener than wrong, but we ought to do so in a predictable ratio of cases.[43]

In order to be able to decide between least squares and threshold of perception, Peirce and Jastrow had to analyse their responses to stimuli with unprecedented fineness. Fechner's experiments with weights consisted in lifting buckets that he filled up himself, and assuming himself to be able to tell in good faith whether he perceived a difference in sensation. Peirce and Jastrow introduced an apparatus that would split Fechner's experimenter into two: an operator providing the stimuli and a subject judging on the sensations. They also split the subject's judgement into a "confidence judgement" and a judgement proper, a feature that we analyze in the next section (Section 4). Thanks to the

41 Quoted in Max Fisch, "A Chronicle of Pragmatism," in *Peirce, Semeiotic, and Pragmatism*, ed. Kenneth Laine Ketner and Christian J.W. Kloesel (Bloomington: Indiana University Press, 1986), 119–120. Emphasis added.

42 Stephen M. Stiegler, *The History of Statistics the Measurement of Uncertainty before 1900* (Cambridge: The Belknap Press of Harvard University Press, 1986), 14–35.

43 Peirce and Jastrow, "On Small Differences," W5, 123.

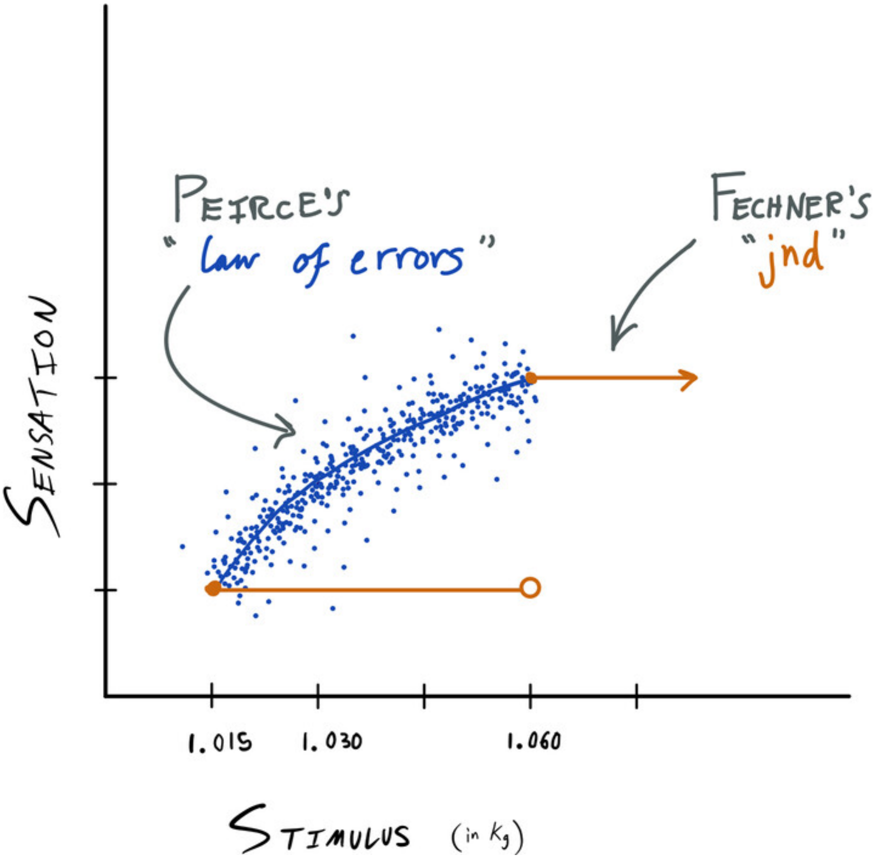

FIGURE 1 Our theoretical reconstruction of how the data would look if either the method of least squares or the notion of "Schwelle" were true of perception; scatter plot

apparatus, the subject was unaware of the weights they would experience (i.e., "blind") and the operator was removed from the position of deciding which weights to apply (randomization). The separation between the subject and the operator was made complete with the use of a screen to hide the one from view of the other, and the randomization effect was accomplished with a deck of cards, which we describe below (Section 4). Finally, to prevent other senses from contributing to the assessment of weight, Peirce and Jastrow decided to utilize a device—a modified Fairbanks scale—which transformed Fechner's exercise of weight lifting into weight sensing, i.e. an exercise in discriminating between different pressure stimuli.

How could a scale be transformed into an instrument to apply pressure? When we ordered a 1912 Fairbanks postal scale from an online antique shop, we hoped that the answer would come with the mail. Unfortunately, the instru-

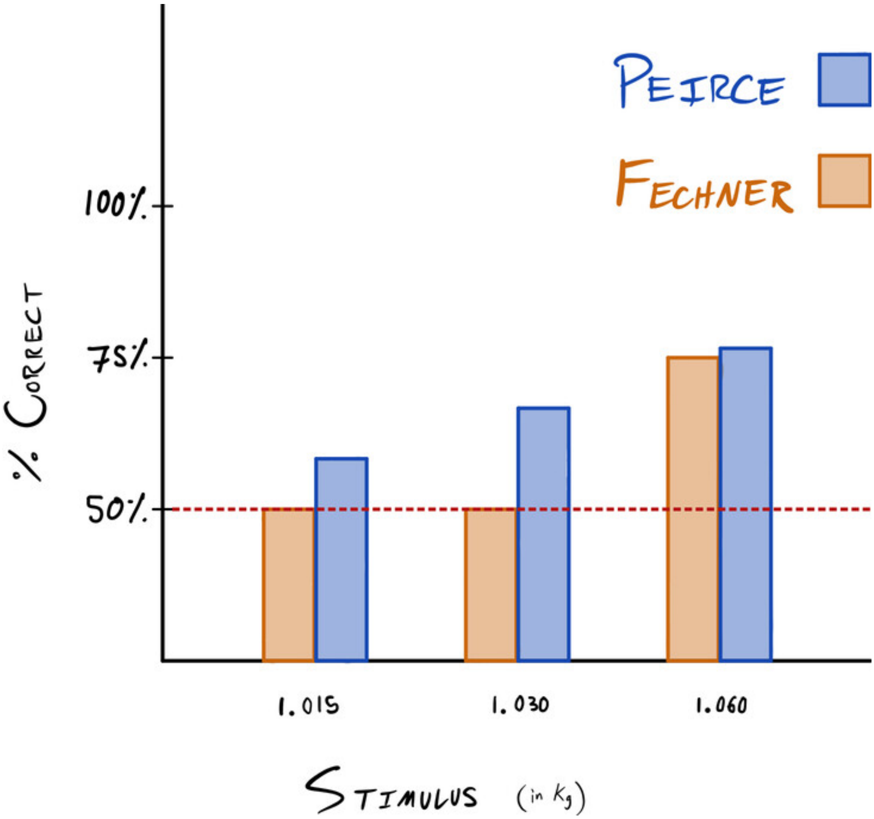

FIGURE 2 Our theoretical reconstruction of how the data would look if either the method of least squares or the notion of "Schwelle" were true of perception; bar chart view

ment itself did not provide any obvious answer to this question. In the description of the apparatus quoted above, Peirce and Jastrow say that "pressure to the finger" was conveyed by "a square enlargement (about one-half inch square)" fixed "upon the end of the beam"; they added that the pressure thus conveyed was one fourth of the weights in the pan, and closed by saying that "the differential pressure was produced by lowering upon the pan of the balance a smaller pan into which the proper weights could be firmly fixed."[44] We had difficulty imagining how pressure could be conveyed by a beam moving upwards when weights were placed in the pan, and initially we thought that the most immediate way to sense pressure would be to place one's finger *directly underneath* the balance's pan. This method however had the disadvantage of producing

44 Peirce and Jastrow, "On Small Differences," W5, 128.

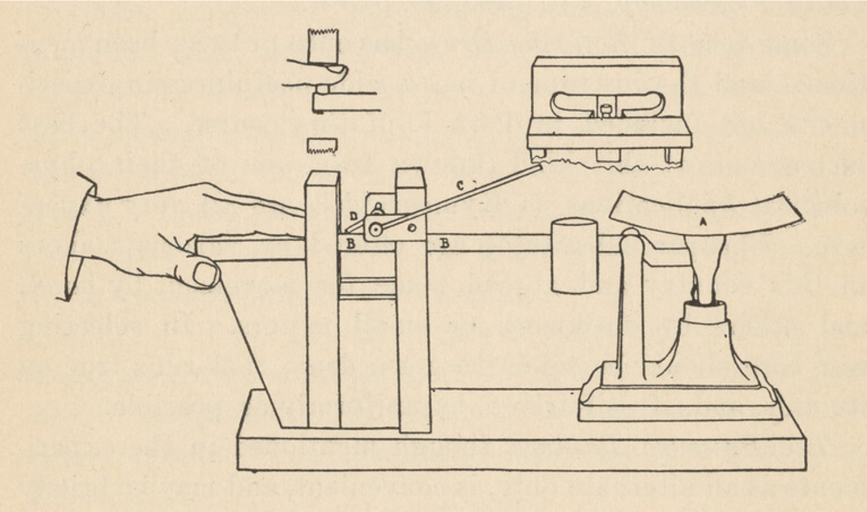

FIGURE 3 Diagram of The Pressure Balance, as found in Sanford (1895–1898), 418. The general description reads: “A is the pan of the scale, B its arm, C the lever, and D the cam.”

too intense, even painful sensations; moreover, subject and operator could not be strictly separated, since they both had to interact with the same part of the balance. While we had not started with the archaeological intention of reproducing exactly the historical instrument Peirce and Jastrow used, our use of the Fairbanks scale was clearly so far off the intended use as to prevent any meaningful investigation into the measuring of sensation and the best way to model it. Prompted by Jastrow’s observation that such instrument was “the forerunner of all the improved pressure-balances since employed”[45] we set out to find “improved” pressure-balances, in the hope that they would afford the key to reverse-engineer the modified Fairbanks scale.

Eventually, we found a description of the pressure balance—complete of illustration (Fig. 3)—in Edmund C. Sanford’s 1895–1898 *A Course in Experimental Psychology*:

> *The Pressure Balance*, though mentioned in the experiments as an alternate only, is convenient, and may be briefly described. *The one in the Clark laboratory, made after the suggestion of Professor Jastrow*, consists of a medium-sized Fairbanks’ letter-scale provided with a wooden base and

45 Jastrow, “Joseph Jastrow,” 136.

FIGURE 4 The reconstructed apparatus

> hand support, and a lever and cam for removing the pressure from the finger. The general construction will be clear from the accompanying cut.[46]

Thanks to the illustration, we were able to start reconstructing the pressure balance that Peirce and Jastrow might have used (Fig. 4); at the same time, we kept searching for remains of the original instrument, but neither Clark University archives, nor the University of Wisconsin (where Jastrow spent most of his career), nor the Jastrow Archive at Duke University could provide any information about the missing scale. The Akron Museum of Psychology, which today hosts the largest collection of historical instruments in the United States, also has no records of the pressure balance.

Only after reconstructing the instrument and performing the experiment we found a *photograph* (Fig. 5) of the instrument: the pressure balance sits on the right-hand side of a table crowded with apparatuses, ready to be exhibited at the World's Columbian Exposition in 1893 (also known as the Chicago World Fair). Since it was Jastrow who "was in charge of the psychological section of the world's fair",[47] the picture provides a much welcome confirmation that our reconstruction of the scale was close enough to the historical model likely utilized by Peirce and Jastrow in 1885.

46 Edmund C. Sanford, *A Course in Experimental Psychology* (Boston: D.C. Heath & Co., 1898), 417.

47 Hull, "Joseph Jastrow," 582.

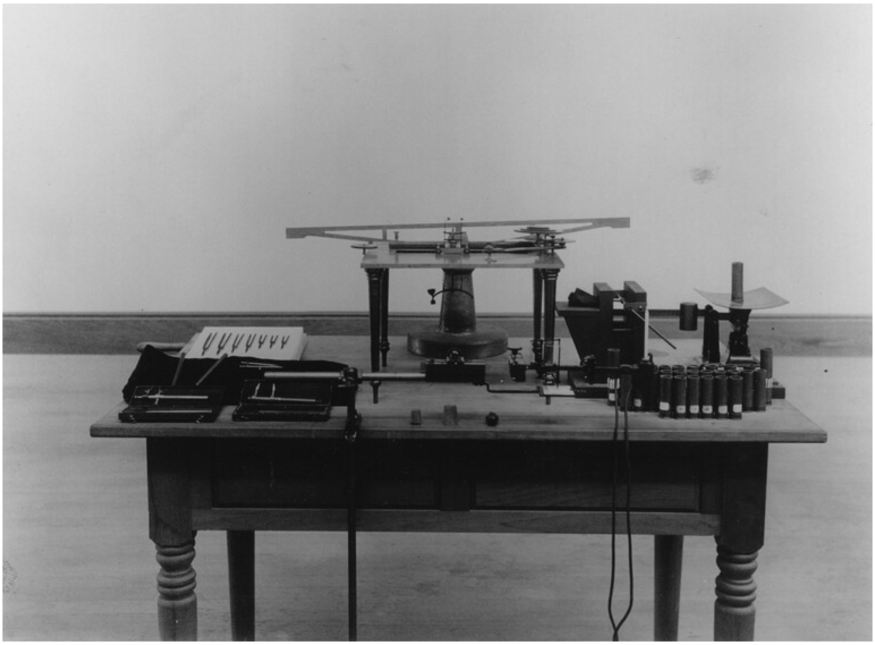

FIGURE 5 "Ensemble of Dermal Apparatus," 1893
PICTURE TAKEN FOR/FROM THE WORLD'S COLUMBIAN EXPOSITION IN 1893. CLARK UNIVERSITY ARCHIVES, WITH PERMISSION; THIS PICTURE WAS PREVIOUSLY PUBLISHED IN JOHN A. POPPLESTON, MARION WHITE MCPHERSON, *AN ILLUSTRATED HISTORY OF AMERICAN PSYCHOLOGY* (AKRON: AKRON UNIVERSITY PRESS, 1994), 24.

In the next section, we describe our experience in operating the modified scale, our results, and how this helped us to better understand Peirce and Jastrow's 1885 experiment and their alternative path to an experimental psychology without fixed "measuring rods."

4 Reproducing the 1885 Experiment

In order to better understand the meaning of C.S. Peirce and Joseph Jastrow's 1885 experiment "On Small Differences in Sensation" and its relevance for problems of measurement in psychology, we undertook three historical replication tasks: (1) the construction of the experimental *equipment* (illustrated above), (2) the reconstitution of the experimental *method*, and (3) the collection, recording, and analysis of experimental *data*. In the following, we focus on (2) and (3). In the previous section, we sketched in broad strokes the purpose of Peirce and Jastrow's experiment and the main function of the apparatus; in what follows, we give a much more detailed account of the experiment's components and its phases, described from the perspective of our own re-enactment of it.

We were able to render the method of experimentation recoverable after many attempts to reconstitute Peirce and Jastrow's procedures. While never aiming at a one-to-one replacement of their historical materials, we realized how delicate the balance between faithful historical reconstruction and philosophical (and scientific) understanding of the experiment was.[48] For example, experiencing pressure on the wrong side of the Fairbanks scale (as discussed in Section 3) was both historically inaccurate and methodologically unsound.

Throughout this project, we realized that we re-enacted Peirce and Jastrow's experiment on two levels: sensing and designing. First, at the level of sensation, we experienced being both the experimenter and the experimental subject, switching off roles regularly, just as Peirce and Jastrow did. Second, at the level of design and redesign, we adjusted some features of the original instrument to our level of carpentry skill and available materials, and limited the number of trials to manage our sense of fatigue within the time that we could dedicate to the project. This longer-term design process too was experiential, informed by the sensory experiences of trying out the protocol. We realized the value of both levels of historical replication—participation and design/calibration—as we came to understand ourselves as part of the experimental apparatus. Alongside the instrument, our perception too was progressively calibrated and refined.

4.1 *Gaining a Sense of the Experimental Protocol: As Subject*

In this section, we go over the basic features of the experimental protocol itself, and what it is like to experience it first-hand. On the pan of a Fairbanks postal scale, a "baseline" weight is placed: 1000 grams (1 kg). The subject's finger rests on a support in such a way that the fingertip will be in contact with the tip of the scale's beam, and prevented from moving upward when pressure is applied (Fig. 6). Any weight which is placed on the scale thus exerts pressure on the tip of the finger (actually corresponding to of 1/4th the weight, as described by Peirce and Jastrow; however, it is not crucial to calculate).[49] A round, spongey

48 The details "behind the scenes" of the first-hand process of discovering how to construct the apparatus, its accompanying accessories, and how to conduct the experiment will be material for a future paper. We thank an anonymous reviewer for pushing us to clarify this point.

49 Because additional weights also only apply one-fourth of their pressure, and the relevant

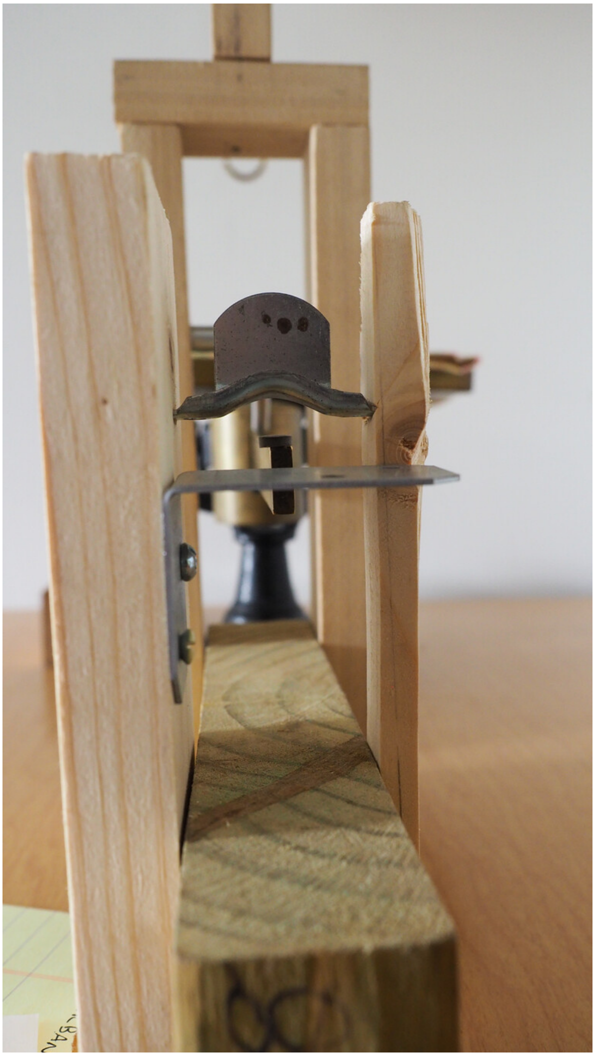

FIGURE 6
The reconstructed apparatus, subject's view

sticker was placed on the end of the beam for comfort and to prevent possible disturbing sensations (noise) coming from the metal's differing temperatures over time.

Each observation starts with the subject verbally communicating that they are ready to the experimenter. As the experimenter releases the beam of the

feature of the experiment is the ratio between the base weight and the base weight + the additional weight, the ratio is unchanged whether we calculate based on the weights themselves or the ratio of pressures which result from the weights.

scale, the subject's fingertip is pressured upwards. Once the subject feels that they have sensed the first stimulus enough (usually for a second or two), the subject calls for a "Change!" to the weights.

Cued by a randomly drawn card, the experimenter will then follow *one* of these two protocols:

– **Increase Protocol:** The baseline weight (1000 g) had been present, and an additional weight is then added (for example, and additional 15 g, for a total of 1015 g);
– **Decrease Protocol:** The baseline weight + the additional weight had been present, and the additional weight is then taken away (leaving behind only the 1000 g baseline weight).

Once the subject feels this change (whichever it may be) to their satisfaction, they again call "Change!". The experimenter then changes the situation of the weights back to what it was originally. The subject again feels the change. When they are ready to pronounce their judgment, they say so; the experimenter brings down the beam, releasing the subject's finger from any pressure. The subject now makes a judgment of whether the weight was first increased (and then decreased again) by saying "increase," or first decreased (and then increased again) by saying "decrease."[50]

Before stating their judgment of "increase" or "decrease," the participant also states their own self-assessment of the confidence they feel in being right about their judgment.[51] They do so by reporting their confidence on a scale of 0–3, with the following agreed upon meanings:

> "0" meaning "I feel as though I am merely guessing;"
> "1" meaning "I feel a slight inclination towards this assessment;"
> "2" meaning "I feel somewhat confident in this assessment;"
> "3" meaning "I feel as confident in this assessment as anyone could hope to be."[52]

50 Peirce and Jastrow's reasoning for the presence of a double change in each observation is the following: "when each experiment embraces a double change, this difference in the amount of sensation caused by an increase and decrease of pressure affects every experiment alike, and the liability to error is constant." Peirce and Jastrow, "On Small Differences," W5, 125.

51 The importance of this passage is discussed in Section 4.5.

52 The meanings which we employed differ in wording only slightly from Peirce and Jastrow's original instructions:

> "0 denoted absence of any preference for one answer over its opposite, so that it seemed nonsensical to answer at all.
> 1 denoted a distinct leaning to one alternative.

The experimenter records the response (see Section 4.3), and one observation is concluded.

4.2 *Gaining a Sense of the Experimental Protocol: As Experimenter*

From the experimenter's side, the primary responsibilities are to enact the experimental protocol with as little disturbance to the subject's sensory experience as possible, while quickly recording their responses. Before each of the sensory trials, the experimenter sets up the weights to perform an "increase" or "decrease" according to cues dictated by two decks of 25 cards (50 total). We used black cards to cue the experimenter to perform a "decrease" and red cards for "increase." Because one set of 25 will have 12 black cards and 13 red, the other vice-versa, we made sure to rotate the two decks, so that for each weight the subject would be exposed to exactly the same amounts of "increases" and "decreases." To randomize the order in which the "increase" and "decrease" trials take place, we shuffled the cards following a method described by Peirce himself in another text.[53]

Between each set of 25 observations, the participant is allowed to take a break to prevent finger (and mental) fatigue. For our purposes, we performed two full rounds of 50 observations on each of three weight increases: 15 g, 30 g, and 60 g.[54] This allowed us to record 300 observations each, totaling 600 observations between us both. Nonetheless, many more observations were necessary to understand how the apparatus worked, to practice making con-

---

2 denoted some little confidence of being right.
3 denoted as strong a confidence as one would have about such sensations" (Peirce and Jastrow 1885, 77).

We felt that by modernising the language, we were able to more closely align with one another's interpretation. For example, "some little confidence" was changed to "somewhat confident," etc., as this is a more familiar phrasing that we felt we could associate more consistently with a mental experience. For our replication purposes, it was more important to us that we understood what we were recording than keeping the original language. We wish to thank an anonymous reviewer for pushing us to clarify this point.

53 In the 1887–1888 manuscript "A Guess at the Riddle," (now published in W6) Peirce describes a method for shuffling cards "with extreme thoroughness": "Cards are almost never shuffled enough to illustrate fairly the principles of probability; but if after being shuffled in any of the usual ways, they are dealt into three packs and taken up again, and then passed from one hand into the other one by one, every other one going to the top and every other one to the bottom of the pack that thus accumulates in the second hand, and finally cut, the shuffling may be considered as sufficient …" (W6, 192, n. 1). We thank André de Tienne for drawing our attention to this passage.

54 While Peirce and Jastrow used other weights at the beginning of their experimentation, the 15–30–60 g weights were the three they settled on for the majority of their experiments.

fidence judgements, and to smoothly and accurately collect data with cards. Only after many trials and adjustments did we feel confident that we were ready to record data for analysis. All in all, we performed this task more than 820 times.

4.3 *Data Recording*

Per Peirce and Jastrow's instructions,[55] we prepared the deck of 50 cue cards for the role of data recording by cutting them into squares and punching holes in three corners, according to the possible levels of confidence that the subject would have: no holes for no confidence, 1 hole for a confidence degree of 1, 2 holes for a confidence degree of 2, and 3 holes for a confidence degree of 3. If the participant's assessment was correct, the cue card would be laid face-up, and if it was incorrect, face-down. The card would also be turned such that the top right corner of the card, whether face-up or face-down, showed the level of confidence. For example, if the assessment was correct, and the participant claimed to have a 0-level of confidence, the card would be placed face-up and turned in such a way that the corner of the card with no holes ("0" confidence) was on the upper right corner. The next judgment was recorded by a new cue card, which was then placed on top of the pile of previous cards, being careful to keep the cards straight in the manner that they were placed. Eventually, this pile was then used to record all observations on a table document.

While this method of recording may seem eccentric when reading Peirce and Jastrow's paper, we came to appreciate how important a quick and easy flow of experimental protocol and data recording is for a sensory experiment. While playing cards seem to be arbitrary, they have distinct advantages. First, unlike other paper cards, they are hardy and meant to be shuffled repeatedly. They also have two binary variables already present in their design (red/black = decrease/increase, front/back = correct/incorrect), and easily accommodate the recording of a four-scaled ordinal variable on their four corners (0, 1, 2, and 3 holes in corners = 0–3 confidence levels). Cutting the cards square allowed for turned cards to fit neatly into the pile, with the upper-right corners (which stood for the confidence-level) flush.

In summary: With a single "flip" of a card, accomplished in a second or two, the experimenter records three variables for future analysis, in the midst of the experiment, with one hand. This is important because the physical and mental fatigue of the participant increases the longer they must keep their hand and their mind ready for observation and judgment. Peirce and Jastrow intended

55 Peirce and Jastrow, "On Small Differences," W5, 130.

to capture the ability of the mind and senses at their best; fatigue muddies the sensory experience, and a tired mind has difficulty assessing its own confidence. For sensory experiments, we came to discover that smooth data collection allows for better, and more, data. With more observations between breaks, one is able to collect more data with less time in between sessions, hopefully leading to more similar data samples. Peirce and Jastrow collected their data over several months,[56] with peak efficiency in data collection allowing them to record "150 experiments on each of us were taken at one sitting of two hours."[57] After many previous practice trials over the course of some months, we were able to accomplish our final batch of 600 observations for analysis within the span of five days, without too much fatigue endangering the quality of our results.

4.4 *Data Analysis*

We conducted 25 trials for each weight in the order Peirce and Jastrow suggested: 60 g, 30 g, 15 g; 15 g, 30 g, 60 g; then repeated, arriving at 300 total trials (per person). After each session of 25 observations was concluded, we took the pile of cards and, one by one, referenced them to record the following data in a spreadsheet: Correct/Incorrect (answer), Degree of Confidence, Red/Black (card), (additional) Weight Value, Person, Date, and Baseline Weight (always 1000 g, for the dataset we analyzed). For the purposes of our analysis, Correct/Incorrect, Degree of Confidence, and Weight Value were the most important variables. We also looked at our data split out by person, in case we saw any significant departures that prohibited aggregating our data together.[58] The details of our analysis is described below.

Prior to any data aggregation or analysis, we sketched out images of the primary analyses we hoped to produce and the patterns of data we expected to see, based on Peirce and Jastrow's experimental claims and what we hoped to learn from our replication. This allowed us to compare our empirical findings with what we anticipated seeing (see Figure 1 and 2 above).

First, we found that our high-level findings matched qualitatively with Peirce's predictions based on the law of errors (compare Figure 7 with Figure 2). The percentage correct was above 50 % (chance) for even a very slight weight increase (15 g, in addition to the baseline weight of 1000 g).

56 Peirce and Jastrow, "On Small Differences," W5, 132.

57 Peirce and Jastrow, "On Small Differences," W5, 130.

58 For reasons of space, we do not elaborate on this aspect here. This is elaborated in a separate paper discussing our process of discovery and analysis in conjunction with the small deviations of our experiment from Peirce and Jastrow's original one.

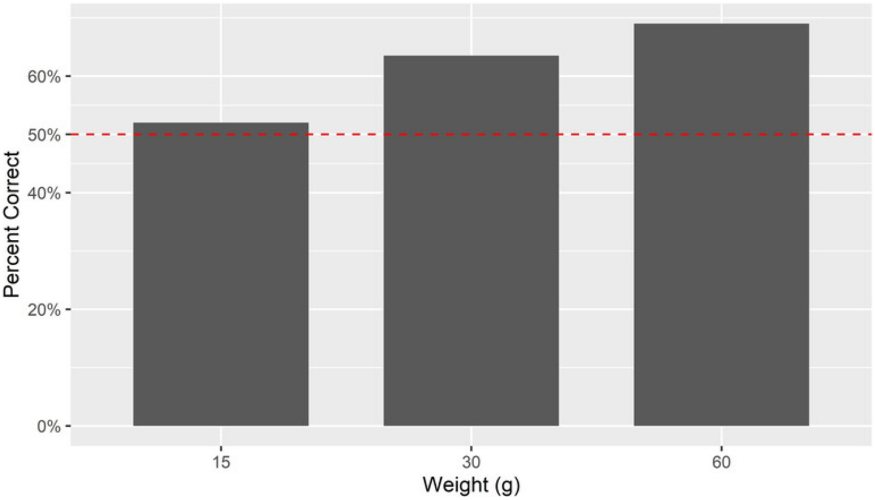

FIGURE 7 Assessments, by weight. Increasingly higher percentages of correct answers were observed for larger weights.

We did not analyze whether our assessments for 15 g were statistically significantly above chance, since our sample is limited (only 200 observations for each weight category). Our interest was in the qualitative trend: the percentage correct gradually increased as weight increased. This pattern corresponds with the prediction based on the law of errors; the bars do not abruptly shift from near 50% to near 100%, as the theory of a "just noticeable difference" would predict (see Figure 2).

From a scatterplot of our raw data (Figure 8), we observed that not only did a higher weight difference gradually lead to greater shares of correct answers, but that the correct answers given had higher levels of confidence. Indeed, the higher the confidence level (the darker shades of blue), the more likely the assessment is correct. This confirmed the role of confidence level in this experiment as meaningfully related to the consciously perceived sensation that was experienced. Yet, while the internal experience of confidence indicated a range of conscious sensation, it is not a perfect picture of the stimulus. There were several observations in which we felt somewhat confident and yet we were wrong (even two very-confident observations which were incorrect). We came to understand that this also is predicted by Peirce and Jastrow's interpretation of the law of errors.

The level of confidence, in addition to not always lining up with the stimulus, does not always line up with the consciously perceived sensation. As seen in the light grey bars in the chart below (Figure 9), when we thought that we were purely guessing (that is, with Confidence of "0"), without a hint of indication

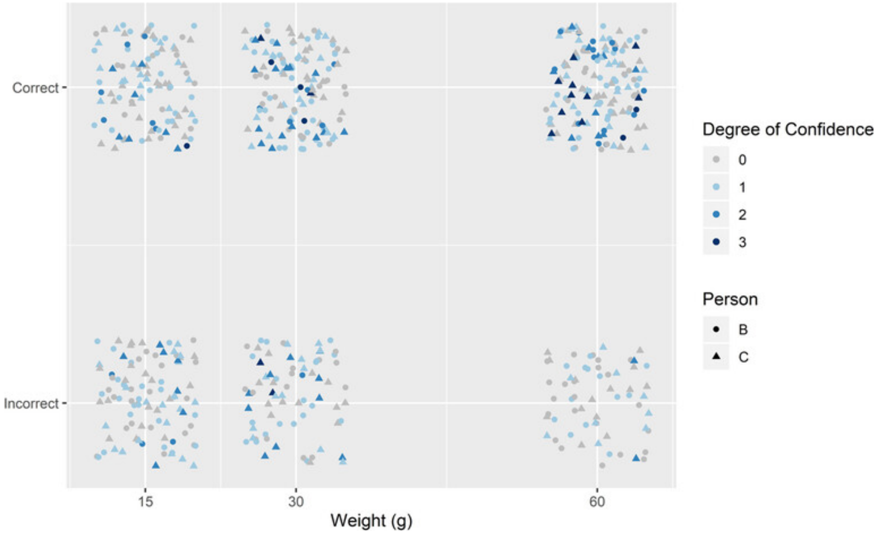

FIGURE 8 Correct/incorrect assessments, by weight. Scatterplot of raw data shows larger weights associated with more correct answers, and correct answers associated with higher degree of confidence. Exceptions can be observed.

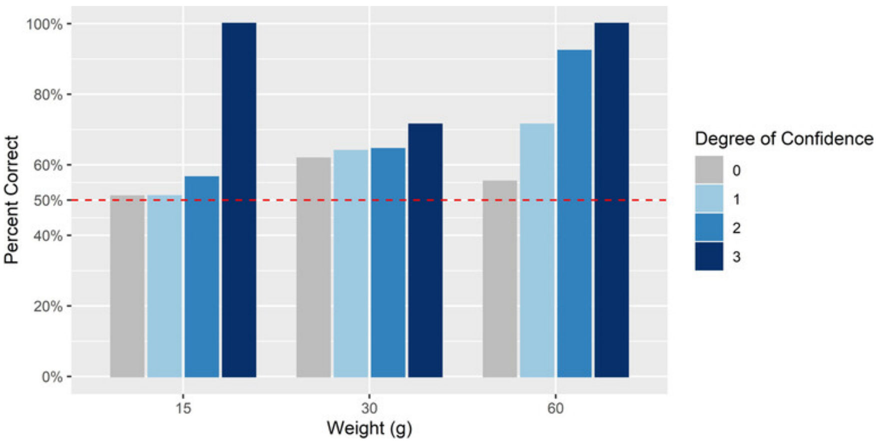

FIGURE 9 Assessments, by weight and degree of confidence. Meaningful relationship between increasing degree of confidence and correct answers, for each weight value (note that only one observation of 15 g weight-increase had degree of Confidence-3.)

either way, we were actually "guessing" well above chance. This phenomenon, of feeling like one is guessing and yet succeeding, is what Peirce and Jastrow emphasize in their results. They sought to show that one could not, as Fechner did, simply assume that when one no longer consciously thinks they feel any difference, that they indeed do not. Peirce and Jastrow showed that even when one feels as if they feel nothing (have 0 Confidence), they actually are unconsciously perceiving. The use of confidence level is one of the true innovations and major contributions of their experimental protocol that differed importantly from Fechner's.

4.5 *Discussion: Confidence Judgment and Its Impact on Experimental Design*

From the experience of being the experimental subject, we could confirm Peirce and Jastrow's surprise about how often one succeeds at arriving at the correct answer, even when one feels they are merely guessing (Confidence-0) or that they barely have any indication of a change one way or another (Confidence-1). Equally surprising are the times when one feels fairly certain (Confidence-2) or even very certain they are correct (Confidence-3), only to find out that what they sensed was in fact incorrect. These phenomena can be observed in our raw data scatterplot (Figure 8).

According to Peirce and Jastrow, this is demonstrative of the difference between sensing and assessing our sensations. We call "primary sensation" the actual sensation, which may be unconscious, and "secondary sensation" our conscious assessment of what we are sensing. Peirce and Jastrow hypothesized that traditional experiments on sensation actually measured our conscious perception only. Thus, the "Differential Threshold of perception" [Unterschiedschwelle] theorized by Fechner would play a role only in measuring secondary sensations: "We are therefore forced to conclude that if there be such a phenomenon, it has its origin, not in the faculty of sensation, but in that of comparing sensations."[59] Yet, how is it possible that our primary sense, even when confidently felt by our secondary sense, could be wrong?[60]

59 Peirce and Jastrow, "On Small Differences," W5, 123.

60 This question has been variously tackled by Peirce scholars, who normally situate it in Peirce's "mature theory of perception" of 1902–1903. For an account of how Peirce may reconcile the irresistible "insistence" of experiential qualities (called "percepts") on our senses and the fact that our consciousness only deals with interpretations of them (the "perceptual judgements") see e.g. Catherine Legg, "Idealism Operationalized. How Peirce's Pragmatism Can Help Explicate and Motivate the Possibly Surprising Idea of Reality as Representational," in *Peirce on Perception and Reasoning: from Icons to Logic*, ed. Kathleen A. Hull and Richard Kenneth Atkins (New York: Routledge, 2017), 13–40.

Peirce and Jastrow's view of human perception was that it itself was a measuring process with a corresponding distribution of error. If perception of pressure is normally distributed, as in accordance with the law of errors, then we should expect that we perceive pressures "above" and "below" the stimulus pressure. On rare occasions (representing the "tails" of the normal distribution) we will perceive a stimulus as much bigger or much smaller than it actually is (recall the spread of blue data points in Figure 8). In that sense, our incorrect answers were correctly stating what we perceived, but our perceptions were mismatched with the stimulus. In this view, the stimulus is refracted through human perception, which only on-average senses the level of stimulus as it is.

5 Conclusion

Peirce and Jastrow's 1885 paper, and the modified Fairbanks scale that they developed for it, has been the object of some attention for its pioneer role in North American psychology and in the history of scientific methodology. It has also been further investigated by Peirce scholars interested in the relationship between Peirce's experimental psychology and his logical practice.[61] However, thus far its relation to the core issues of early experimental psychology, which mostly concern the possibility of measuring sensation, has received less attention. Its dense exposition, the presentation of data in tables (rather than in graphical form), and the absence of any illustration of the instrument used have further impaired the reception of the paper among historians and philosophers of science.

Peirce and Jastrow proposed that sensation may be described using the same mathematical theory used to describe observations in astronomy, namely the law of errors, as employed in the method of least squares. In doing so, they are aligned with a long philosophical tradition which sees the human senses as (more or less accurate) measuring devices. In this framework, problems of measurement traditionally attributed to the physical sciences naturally apply to the measurement of sensation. The relevance of measurement for the scientific status of psychology both before and after Peirce and Jastrow is described in Section 2.

Measurement needs instruments, and so experimental psychology needed apparatuses. Section 3 described in broad strokes the purpose of Peirce and Jas-

61 See Cadwallader, "The First American Experimental Psychologist;" Hacking, "Telepathy;" Cristalli, "Experimental Psychology."

trow's 1885 experiment and then told the story of our material engagement with it, through the reconstruction of the "modified Fairbanks scale." We highlighted how our quest for a more historically accurate instrument was motivated by the impossibility to reproduce the observations described with a simple (unmodified) historical postal scale. The necessity to find models for Peirce and Jastrow's scale led us to put together historical documents which have so far been disregarded by historians of measurement and of pragmatism alike. We documented the process of reconstruction with pictures coming from our sources as well as from our reconstructed "pressure balance."

Finally, in Section 4 we described in detail our re-enactment of Peirce and Jastrow's experiment, as well as the main insights that this experience provided on the 1885 paper. By viewing the senses as yet another measuring device, Peirce and Jastrow were able to apply the same measuring protocols (method of least squares, theory of errors) already developed in the physical sciences for the observation of stars. In order to be able to separate our conscious judgment of perception from what we actually sense, Peirce and Jastrow modified Fechner's original experimental design in many ways: they constructed a setting involving two people, so that the subject would actually not know to which pressure difference they were being exposed to; they introduced devices (such as a screen) between subject and experimenter to prevent unintended cuing; and, most importantly, they introduced a "confidence judgement" alongside the actual judgement of sensation.

To reproduce Peirce and Jastrow's experiment, we had to learn unfamiliar tasks of sensory discrimination and how to accomplish quick and smooth data recording. In the process, we came to appreciate the ingenuity of their method, such as the use of punched cards to "randomize" the trials while at the same time recording our data. Moreover, the search for Peirce and Jastrow's instrument led us to unearth its legacy beyond the 1885 experiment, to which its inclusion in the World's Columbian Exposition of 1893 and in Sanford's 1898 manual of experimental psychology testify. These insights demonstrate the importance of experimental reconstruction in revealing a more accurate understanding of the written text, and the potential for exploratory hands-on research to play a complementary role in historical inquiry.

Acknowledgments

While the work behind this paper is fully collaborative, Cristalli is the main author of sections 1–3 and Jackson of section 4. Please address questions on Peirce, Jastrow, or the material reconstruction of the psychological apparatus to Cristalli, and questions on our re-enactment of the experiment and on the analysis of our findings to Jackson.

We would like to thank Jutta Schickore and William Newman for their infectious love for history of methodology and empirical methods of historical inquiry, as well as for drawing many scholars like us together in Bloomington, Indiana for the Sawyer Seminar, “Rigor: Experimental Control, Analysis, and Synthesis in Historical and Systematic Perspectives,” a year-long research institute hosted by the Department of History and Philosophy of Science and Medicine at Indiana University—Bloomington and funded by the USA-based Mellon Foundation. Additionally, we owe our sincerest gratitude to André De Tienne, Director and General Editor of the Peirce Edition Project, whose knowledge of Peirce and his writings is only matched by his joy in aiding scholars in any curious historical inquiry. We also would like to thank Brad Luen, who graciously aided our understanding of the statistical analyses in our source material, in addition to serving as an experimental subject as we developed the instrumental set-up. Lastly, we owe appreciation to the citizens of Middle Earth Co-operative Housing, who put up with several weeks of us woodworking on their porch while arguing about 19th-century philosophy of measurement.

During the time of researching and writing this piece, Claudia Cristalli was supported by a Mellon Grant postdoctoral research fund within the project “Rigor: Experimental Control, Analysis and Synthesis” at Bloomington, Indiana. Rebecca Jackson was supported by the John C. Slater Predoctoral Fellowship in the History of Science, Technology, and Medicine (2021–2022) from the American Philosophical Society, which graciously allowed her to travel during her residential fellowship to conduct the empirical portion of this work.