How Did Early North American Clinical Psychologists Get Their First Personality Test? Carl Gustav Jung, The Zurich School of Psychiatry and The Development of the "Word Ass....

Article *in* History of Psychology · May 2022

DOI: 10.1037/hop0000218

CITATIONS
4

READS
780

1 author:

Catriel Fierro
Autonomous University of Barcelona

55 PUBLICATIONS **185** CITATIONS

SEE PROFILE

History of Psychology

ISSN: 1093-4510

2022, Vol. 25, No. 4, 295–321
https://doi.org/10.1037/hop0000218

How Did Early North American Clinical Psychologists Get Their First Personality Test? Carl Gustav Jung, the Zurich School of Psychiatry, and the Development of the "Word Association Test" (1898–1909)

Catriel Fierro

Institute of Humanities and Social Sciences (INHUS), National Scientific and Technical Research Council, Mar del Plata, Argentina

Faculty of Psychology, National University of Mar del Plata

Clinical psychology emerged in the United States during the first decades of the 20th century. Although they focused on intelligence tests, starting around 1905 certain clinical psychologists pursued personality assessment through a specific, nonintellectual kind of test: the word association test as devised by Swiss psychiatrist Carl Gustav Jung (1875–1961) at the Burghölzli psychiatric clinic in Zurich. The test was a key device in the professionalization of North American psychiatry and psychology during the early 20th century: from 1905 onward it was acknowledged, discussed, and applied by experimental and clinical psychologists. However, Jung's original experiments and the development of the test itself have received only superficial or casual attention by historians of science. This article attempts to provide a critical, streamlined, and detailed account on the origin, development, and substance of the Zurich word association experiments. By drawing on heretofore overlooked primary sources, I offer a new, critical perspective on the emergence and development of Jung's test while engaging with its main theoretical and methodological aspects. I show that the test was neither Jung's sole creation nor did it consist of a simple, straightforward set of tasks. Contrarily, it was the result of a highly collaborative, multilayered institutionalized research program on linguistic and mental associations. The program, its data and its assumptions fueled several debates and data-driven discussions at Zurich, precluding the test from achieving a stable, standardized character. As a result, the history of Jung's program reflects both the advances and the limitations of early 20th-century personality testing.

This article was published Online First May 5, 2022.

Catriel Fierro https://orcid.org/0000-0001-7635-7012

This work is part of the research project "Statistics, experiments and the clinical method: The professionalization of clinical psychology in the United States and Argentina, 1920–1970," funded by Argentina's National Scientific and Technical Research Council. A previous version of this work was presented as a paper at the 53rd Annual (Virtual) Meeting of Cheiron, June 15–17, 2021. I thank Ian Davidson, David Devonis, David Schmit, Larry Stern, Michael Sokal, and Jacy Young for their helpful comments on my article. I also wish to thank Nellie Thompson for kindly providing me with relevant bibliography on the issue. Finally, I also thank Christopher Green for encouraging me to submit the article to this journal. I am nonetheless responsible for any of the article's limitations.

Correspondence concerning this article should be addressed to Catriel Fierro, Faculty of Psychology, National University of Mar del Plata, Deán Funes 3250, Fifth Building, Third Floor, "Secretaría de Investigación y Posgrado," Buenos Aires B7600, Argentina. Email: catriel.fierro@gmail.com

Keywords: history of clinical psychology, personality testing, experimental psychopathology, word association, European psychiatry

Clinical psychology emerged as a specialized field in the United States during the first decades of the 20th century. As researchers in "abnormal" mind states and behaviors, clinical psychologists quickly became associated with psychological tests. Indeed, during the 1910s the diagnostic technique that dominated the psychological field was the Binet-Simon scale, including its revisions and derivatives (Reisman, 1991; Zenderland, 2001). Nonetheless, starting around 1909 psychologists conducted personality assessment through a specific, nonintellectual kind of test: the word association test (WAT) as devised by Swiss psychiatrist Carl Gustav Jung (1875–1961) at the University of Zurich's "Burghölzli" clinic.[1]

Jung's test was first introduced to the United States through a series of enthusiastic reviews by North American psychiatrists (Meyer, 1905; see also Taylor, 1998). As early as 1906 it was defined by Adolf Meyer (1866–1950), a former Burghölzli alumnus, as "so far the nearest approach of an experimental test to the combination of a qualitative and quantitative inquiry into the stream of mental activity and its most frequent disturbers" (Meyer, 1906, p. 280). Indeed, Jung's approach was defined as the *essential link* between "the excessively quantitative tendencies of a great deal of laboratory psychology and the excessive tendency to measure symptoms in mental disease merely by the anomaly or absurdity of the content" (Meyer, 1905, p. 250; see also Leys, 1981, 1985). This appreciation also boosted Jung's reputation among psychologists: during the spring of 1909 G. Stanley Hall and his assistant Amy Tanner used Jung's test to assess the authenticity of the spiritualistic phenomena channeled by illustrious medium Mrs. Piper—to all effects a clinical, diagnostic use of the test (Rosenzweig, 1992, pp. 86–89). Hall found the method sound and the same year he invited Jung to give a series of lectures in the context of Clark University's 20th anniversary in September 1909—lectures that revolved around the association method (Evans & Koelsch, 1985; Skues, 2012, pp. 68–73; see also Green, 2019, pp. 312–313). From 1909 onward, several empirical and theoretical studies on Jung's method appeared in leading North American psychological journals, especially in Hall's own *American Journal of Psychology*; by the 1910s clinical studies on word association had crossed professional borders, and Jung's method was described by clinical psychologists as "epoch-making" (Kohs, 1914, p. 551; see also Zenderland, 2001, pp. 304–306).[2] Jung's word association experiments (hereafter WAE or WAEs) provided psychology and psychiatry with objective tools for probing the "abnormal" mind.

Given their popularity and their ultimate historical significance, it is striking that Jung's original experiments have received only superficial or casual attention by historians of science. As Homans (1979) and Taylor (1996) have shown, many histories of psychology or psychoanalysis have focused mainly on Jung's latter developments on archetypes and the collective unconscious while omitting any specific analysis of Jung's experimental works. Those historical accounts that have touched upon the issue have not engaged with Jung's early publications, merely acknowledging that Jung worked on word association experiments (Cohen, 2014; Jehle-Wildberger, 2020; Kirsch, 2004; Monahan, 2009). But how did these experiments begin? How did they develop? And who carried them out? Indeed, by avoiding these

[1] According to Baker and Benjamin (2014, pp. 71–72), Jung's was the first personality test used by clinical psychologists. See also Reisman (1991, pp. 96–98).

[2] For a brief analysis of the reception of Jung's test in the United States, see Green and Feinerer (2016). Davidson (2020) offers a broader analysis.

questions historians have tended to treat the experiments as self-evident and self-explanatory (for histories of psychoanalysis, see Clay, 2016, pp. 80–87; Makari, 2008, pp. 188–197; Meyer, 2013, Zaretsky, 2004; for histories of psychology, see Rosenzweig, 1992, p. 136; Smith, 2013; Walsh et al., 2014). As a result, historical scholarship has incurred several oversimplifications or misrepresentations. First, it is usually assumed that both the WAEs and the WAT consisted of a relatively straightforward task or series of tasks. The test is usually depicted as having involved an experimenter calling a "standard list" of 100 stimulus words (hereafter SW or SWs) to a subject who is requested to say the first response word that comes to mind (Taylor, 1998, p. 102; see also Bernet, 2013, p. 255, Homans, 1979, p. 45; Meyer, 2013, pp. 205–206). The experimenter then measured each reaction time (RT), sometimes also classifying the resulting responses according to their content (Clay, 2016, pp. 83–84; Homans, 1979, p. 45; for similar depictions, see Leys, 1985, pp. 345–346; Shamdasani, 1998, pp. 120–122). Second, the WAT has been portrayed as a single, stable entity that produced transparent experimental data: in Cohen's terms, "the test came to have a standardized form [. . .] and technique" (Cohen, 2014, p. 4). In other words, the results Jung obtained through the experiments are usually seen as conclusive and definitive: no mention is made to either contradictory conclusions or to internal debates at the Burghölzli that could have led to methodological and procedural revisions of the test. Third, Jung is usually depicted as either working alone or assisted by Franz Riklin (1878–1938). As a result, Jung's name and the test have become synonyms (Borch-Jacobsen & Shamdasani, 2013, pp. 58–61; Frey-Rohn, 1990, pp. 3–35; Makari, 2008, p. 256; Shamdasani, 2004, 2011; for similar accounts see Ellenberger, 1970/1994, p. 668, pp. 691–692; Falzeder, 2015, pp. 117–118; Graf-Nold, 2001, p. 86; Skues, 2012, pp. 68–73).

In the context of these assumptions, this article attempts to provide a critical, streamlined, and detailed account on the origin, development, and substance of the Zurich word association experiments conducted at the Burghölzli during Jung's tenure there from 1900 to 1909. By drawing on heretofore overlooked primary sources, I attempt to offer a new, critical perspective on the emergence and development of Jung's WAT while engaging with its main theoretical and methodological aspects. A careful, detailed analysis of the publications by the Zurich school suggests that the WAT was neither Jung's sole creation nor did it consist of a simple, straightforward set of tasks. On the contrary, it was the result of a collaborative, multilayered institutionalized research program on mental association that fueled several debates and data-driven discussions at the Burghölzli. As a result of these debates, the test itself never achieved a stable, standardized character. Thus, its history sheds light on both the advances and the limitations of early 20th-century personality testing.

Unearthing Emotional Complexes: Bleuler, Jung, and Experimental Psychopathology at the Burghölzli Clinic (1900–1903)

Two weeks after obtaining his medical degree from the University of Basel, Carl Jung arrived in Zurich and took up a post as second assistant physician at the Burghölzli Clinic on December 11, 1900. Depicted in Figure 1, the Burghölzli was a 33-hectares cantonal facility equivalent to an American state hospital. Situated in southeastern Zurich, since 1898 it had been under the direction of Eugen Bleuler (1857–1939), a young and innovative Swiss psychiatrist (Ellenberger, 1970/1994, p. 666; Möller et al., 2002; Shamdasani, 2004, pp. 44–45). Indeed, by 1900 the Clinic was rapidly gaining recognition as a leading institution in the study and treatment of severe mental disorders. In marked contrast to other psychiatric clinics focused on the description and classification of patients, early on the Burghölzli adopted a

Figure 1

The Swizz Institution for Epileptics (Middle Left) and the Burghölzli (Upper Right) in 1918

Note. *Source:* Walter Mittelholzer, ETH-Bibliothek Zürch, Bildarchiv/Stiftung Luftbild Schweiz (LBS_MH01-001589). Image in Public domain (https://doi.org/10.3932/ethz-a-000342310). See the online article for the color version of this figure.

"dynamic" outlook which emphasized innovative psychotherapeutic treatments such as hypnosis and suggestion alongside the psychological exploration of behavioral abnormalities (Bernet, 2013, pp. 84–85; Loewenberg, 1995). Bleuler built a staff that was open to dynamic psychiatry, introduced them to new experimental research techniques, and promoted horizontal debates between directors, chief physicians and assistants. By the early 1900s the Burghölzli "was *the* place in the world to go for any young, ambitious, open-minded psychiatrist" such as Jung (Falzeder, 2015).[3]

By the time Jung arrived at the Burghölzli, the clinic was a huge institution with scarce human resources: during the early 1900s, each physician was in charge of around 95 patients (Minder, 2001a). Such scarcity stimulated intellectual exchange and intensified the staff's engagement with clinical issues. Working with particular zeal, Jung quickly ascended the Burghölzli's professional ladder: by 1905 he was head of the Clinic's outpatient service and Bleuler's *Sekundärarzt*—the equivalent to the clinical director of American mental institutions, or a *senior* physician in contemporary terms.[4]

Besides populating the Clinic with young and able psychiatrists, Bleuler introduced them to new experimental techniques for studying psychopathology. Bleuler was especially interested in the role of mental associations in the origin and development of psychopathologies. To Bleuler, every psychical activity rested "upon the interchange of the material derived from

[3] Bleuler practically handpicked Jung by offering the young physician a position as assistant physician in mid-1900, before the latter had even finished his studies (Wieser, 2001, p. 18).

[4] Available chronologies about Jung's changing appointments are often contradictory. An accurate picture can be reconstructed from the original data published in Minder (2001b, pp. 56–57) and Graf-Nold (2001, pp. 82–83).

sensation and from memory traces, upon associations" (Bleuler, 1904/1918b, p. 1). As a psychic activity, associating involved the linking of current processes with memory traces. These associations were commanded by an affective process Bleuler called attention: a dynamic, energy-based type of psychological *adjustment* that operated by facilitating certain definite associations while inhibiting others. Bleuler equated attention to "an electric storehouse" whose function was to serve as a place of contacts and connections "to and from definite directions" (Bleuler, 1905/1918a, p. 283). Those psychic elements that were *attended* entered in associations with the ego, or "ego-complex": that is, with "the presentations, sensations, desires, which at any given moment compose our personality" (Bleuler, 1905/1918a, pp. 281–282). Elements that were *not* having any transaction with the ego were "automatic" and, thus, unconscious.

Bleuler posed that attention was a regular process in a literal sense: each mental association was an effect of the law of association. This law operated by "pairing" elements according to a series of characteristics, such as similarity, contrast, simultaneity, and relationship, in the context of the subject's "whole personality with its inner and outer past" (Bleuler, 1904/1918b, p. 4). Bleuler conjectured that it was precisely this law that was altered in abnormal mental states: associations were weakened or disturbed in psychiatric cases, leading to chronic symptomatic presentations. Nonetheless, by 1900 Bleuler was in a tight spot in regards to the evidence required to back up his claims. On the one hand, he advocated for his theory on association, arguing that there was no better testimony to the value of a psychological theory "than its applicability in psychopathology" (Bleuler, 1905/1918a, p. 268). On the other hand, he admitted that available evidence was inconclusive: while his decades-long clinical experience was not irrelevant, its subjective, unsystematic, and possibly biased character made it unsuitable as experimental proof. As Bleuler himself admitted, he was "not in a position to give a strict proof of the correctness" of his hypotheses (Bleuler, 1905/1918a, pp. 294–295).

Bleuler first introduced WAEs at the Burghölzli around 1900, to submit his ideas to experimental analysis and to produce systematic evidence (Jung, 1905/1973d, p. 320). To the author, the significance of the association experiments rested in the fact that acoustic reactions could stimulate and "revivify" associations, shedding light on the workings of the individual's mind. His first attempt at the experiments involved an initial list of 156 stimulus words, which he used to obtain associations from several patients at the institution. However, Bleuler soon found a methodological obstacle: the lack of statistical, normal standards. In other words, there were no means for differentiating and comparing the associations made by abnormal individuals with those made by well-adjusted subjects. As such, Bleuler's initial results provided no information regarding either the variations of associations in normal persons or the laws behind the "purely haphazard nature" of the associations. To use the experiments as a diagnostic device, standards were needed to differentiate "with statistical precision the associations of the abnormal from the normal" (Jung & Riklin, 1904/1918, p. 8).

Enter the scene Franz Riklin. Born in St. Gallen, Switzerland, Riklin was trained in psychiatry and had arrived at the Burghölzli in 1900 to work as an *Unterassistenten* (an underassistant). Shortly after he had arrived, and before the year was through, Bleuler sent the 22-year-old assistant to Heidelberg for several months to study with Emil Kraepelin (1856–1926) and Gustav Aschaffenburg (1866–1944), two German authorities in the field who championed the use of association experiments for descriptive and classificatory purposes (Bernet, 2013, pp. 188–189; Kerr, 1994, p. 44). Riklin returned to the Clinic in the Spring of 1901; by August, Bleuler had promoted him to assistant physician and had

introduced him to the newly arrived Jung for them to apply the tests to the Burghölzli's residents.[5] Riklin quickly became one of the initial cornerstones of the program, and its *actual* direction pivoted between the two men. For instance, while Jung left for France from October 1902 to February 1903, his colleague was left in charge of coordinating and conducting the experiments. In all, Riklin worked with Jung from 1901 up to late 1904, when he left the Clinic for the Rheinau Asylum.

Jung and Riklin took over Bleuler's project around late 1901. By applying their chief's theory of association to an experimental framework, they entered headlong into the field of experimental psychopathology. For Jung, experimentation was the most valid form of psychological knowledge: the ultimate aim of scientific exploration was the finding of universal laws (Jung, 1973b). Reality was made up of an "infinite number of chance events," which scientists had to wade through looking for general regularities (Jung, 1905/1973d, p. 332). Because laws implied necessity, experimenters had to establish necessary relations between certain stimuli and specific reactions and test them through empirical checks.

Given the (assumed) fundamental role played by associations in both the normal and the abnormal mind, Jung argued that scientific psychological research had to focus on psychical associations as revealed by WAEs. Such research would discover the psychopathological regularities that were needed to ground experimental psychopathology on empirical terrains (Jung, 1903/1973f, p. 410). Drawing from several authors such as Pierre Janet (1859–1947), Theodor Ziehen (1862–1950), and from Bleuler's theory, Jung argued that a psychical connection or association was always established between two psychical elements or processes. "Affect-toned" or "feeling-toned complexes" were examples of such psychical elements. The "affect-toned complex" was defined by Jung and Riklin "the total number of presentations relating to a definite experience that is charged with emotion" (Jung & Riklin, 1904/1918, p. 66). Complexes were definite groups of ideas "cemented" by definite affective or emotional tones. They were "normal" in the sense they were present in every individual. However, they varied in regards to the intensity of their emotional charge: the stronger the complex, the more vivid the inferred emotional tone (Jung, 1905/1973d). Moreover, in individual psyches certain complexes became *too* strongly toned, threatening the psyche's integrity. As a result, they were repressed, in the sense that the ego broke any association with them.[6] However, displaced and repressed complexes exerted a pathological attraction of attention away from the individual's consciousness, which in turn influenced associations. As a result, the subject reacted to external stimuli not arbitrarily, but by deriving reactions from the repressed complex.

In Jung's own terms, association experiments were conceived as "the means of studying experimentally the behavior of the complex" (Jung, 1911/1973c, p. 600).[7] From the clinician's external point of view, psychical associations, complexes, and constellations were represented only by the subject's outer verbal signs: this is, by his or her utterances. Consequently, if the subject was provided with a carefully chosen word and gave a verbal report about his internal reaction, an analysis of the association between the two occurrences

[5] Thus, it was not Jung but Riklin who actually kickstarted the association experiments at the Clinic. Moreover, Jung learned the test from Riklin, not the other way around (Wieser, 2001, p. 21).

[6] Jung defined the ego as "psychologically nothing but a complex of imaginings held together and fixed by the coenesthetic impressions [and by] intentions or innervations" (Jung, 1911/1973c, p. 601). This was basically Bleuler's definition of the ego.

[7] It has been claimed that while Jung worked at the Burghölzli, "the main research focus [at the Clinic] was human memory" (Davidson, 2020, p. 26). However, ever since the inception of Bleuler's research program on word association in 1900, both he and his assistants were studying not only memory but affectivity, attention, emotivity, unconscious representations, dreams, and even general emotional adjustment.

could be carried out to investigate both the association law (experimental psychopathology) and the quality and dynamics of psychological complexes (individual, clinical psychology). With these ideas in mind, around 1901 Jung and his colleagues started studying the influence of attention on the association process.

A Psychological Snapshot of the Mind: Initial Testing on Normal Individuals at the Burghölzli (1901–1904)

Jung and Riklin first focused on normal individuals to establish criteria for comparing clinical cases. Although they admitted that what qualified as an "average normal" was a delicate and relative question, they defined "normal" as the absence of marked psychological eccentricities.

The Zurich school's first publication describing both the experiments and its initial findings was a 200-page paper written by Jung and Riklin and published in the *Journal für Psychologie und Neurologie* from 1904 to 1905 (Jung & Riklin, 1904/1918).[8] They administered a list of 400 words to a sample of 38 Burghölzli employees ranging from 20 to 50 years of age: the paper itself reported a statistical analysis of the associations collected paired with in-depth clinical examinations and interpretations for each case.

Jung and Riklin assumed that difficult and rare words would cause mishearing or prolonged reactions. They chose words that were used in daily life to prevent interferences that would hinder the validity of experimental results. Thus, the first-ever list of SWs used in an association experiment for clinical ends at the Burghölzli (or the second, if we consider Bleuler's failed initial effort) was composed of 231 nouns, 82 verbs, 69 adjectives, and 18 adverbs and numerals. The original procedure went as follows: first, the experimenter explained the task to the subject. Once the experimenter confirmed the subject had understood that he or she had to answer as soon as possible with the first word that came to his or her mind, the experiment started. The word list was divided into three series as to allow for different experimental conditions of distraction of attention. In the first experimental series, 200 words were administered *individually* to each of the subjects, with no change in the experimental conditions. A second series consisted of 100 further reactions carried out during a condition of inner distraction: attending to an acoustic stimulus.[9] Finally, a third series consisted of 100 further reactions carried out during a condition of outer distraction: associating while making pencil strokes of one centimeter in length to the beat of a metronome. After each reaction, the experimenter noted down the reaction word.

Each subject produced 300 to 400 associations; by the end of the experiments our authors had 12.400 associations to analyze. To derive normal statistical figures, Jung and Riklin first attempted to analyze the entire mass of associations. However, the task was near impossible, given there were thousands of possible kinds of combinations made by the 38 subjects. To find a "foothold in the wild chaos," some structure should be imposed on the innumerable individual reactions. As a solution, they resorted to a classification system (Jung, 1903/1973f, p. 412). The scheme was admittedly provisory and imperfect: the ideal classification had to "evolve from the inner psychical data," not from the "association external appearances" or from "logical principles" (Jung & Riklin, 1904/1918, p. 14). However, given the exploratory

[8] Several accounts have erroneously dated this as 1906 (Rosenzweig, 1992, pp. 135–136; Smith, 2013, p. 196).

[9] To confirm the compliance of the subjects toward the distracting stimulus, Jung and Riklin (1918, p. 12) requested each of their subjects an introspective report for each of the produced associations.

nature of the work, Jung and Riklin offered a simple classification based precisely on logical principles.[10] The authors posited the existence of 21 different kinds of associations grouped in four major categories. *Inner* associations involved affinity or similarity between the SW and the reaction, or a subject-object relationship between them. *Outer* associations involved linking through contiguity. *Clang* reactions were prompted by the sound of the SW. Finally, a *residual group* gathered the remaining kinds of associations.

Several findings were observed in enough cases as to guarantee an inductive generalization. First, when attention waned subjects produced fewer inner associations and more outer associations and clang-reactions. Thus, when attention decreased behavior changed "in the direction of habit and routine, that is, of mechanically simple or verbal connexions [sic]" (Jung & Riklin, 1904/1918, p. 122).[11] Because psychological complexes drew attention from the ego, then a "high" ratio of outer to internal associations suggested a dissociated state of attention caused by a complex. Second, because clang reactions were deemed as the "most primitive of similarity associations," their preponderance was also an indicator of a psychological complex drawing from the subject's associative energy (Jung & Riklin, 1904/1918, p. 42).

When summing up their results, Jung and Riklin argued that associations varied in normal subjects under the influence of three factors: attention, education and "the individual peculiarity of the subject" (Jung & Riklin, 1904/1918, p. 167). Although listed third, there is virtually no topic of discussion in Jung and Riklin's article which they do not moderate by invoking individual differences. During the experiments, psychological dissociations were provoked by distracting stimuli; however, in real life the origin of the attention disturbance was very diverse, "if not specific for each individual process": motor excitation, the declining of kinesthetic sensations and the rising of the muscular stimulus-threshold varied greatly among the subjects (Jung & Riklin, 1904/1918, p. 49). Moreover, the *nature* and *content* of the subject's reactions were very diverse, tied as they were to the "individual temperament" of each subject (Jung & Riklin, 1904/1918, p. 132). "Temperament" meant differences in intellectual and dissociative dispositions, leading apparently similar people to produce very diverse associations. As such, the most important associative variations were conditioned by those differences (Jung & Riklin, 1904/1918, p. 168). In fact, Jung and Riklin devoted around 50% of their article to painstakingly detailed, quantitative and qualitative descriptions of the associations carried out by *each* of the 38 experimental subjects. "Dealing with averages in such delicate quantitative relationships is a somewhat daring undertaking," they admitted, so they focused on profiling and exploring each case (Jung & Riklin, 1904/1918, p. 125).

Between 1903 and 1904 Jung and Riklin also attempted to compare the associations of the individuals of their sample who belonged to the same family, suggesting that there were strong resemblances between mothers and daughters. However, they ultimately concluded that their material was not sufficient to explain their observation: "The complete proof and explanation of these apparent facts must await an investigation from material specially collected for this purpose" (Jung & Riklin, 1904/1918, p. 57). Such an investigation was in fact being carried out by Emma Fürst (1875–1939), one of the first women to be trained as a psychoanalyst in Europe, and the only woman to hold an assistantship at the Burghölzli before the First World War.

[10] The authors stated that other, more refined proposals would follow when they had found themselves "in a position to deduce empirical laws from the mental associations" (Jung & Riklin, 1904/1918, p. 12). However, they never produced a second classification.

[11] As an explanation, the authors argued that association was ultimately characterized by muscular tensions: a state that "supplies the psychophysical basis to the accentuated [psychical] complex" (Jung & Riklin, 1904/1918, p. 123). This "somatic connection" decreased when attention was interrupted by either inner or outer stimuli.

Born in 1875 in Bassersdorf, Zurich, Fürst studied medicine at the University of Zurich from 1897 to late 1903.[12] She then moved to Bern to continue with her studies and in 1905 she became an assistant physician at the Burghölzli, being described by a colleague as "one of the most competent members" of the Zurich psychiatrists (Karl Abraham, 1924; quoted in Falzeder, 2002, p. 1027). She carried out her doctoral research under Jung, her own experiments beginning "at the instigation of Professor Bleuler and Dr. Jung [. . .] in 1903 (Fürst, 1910/1918, pp. 407–408). For her dissertation Fürst collected associations from 100 individuals between 9 and 89 years of age who belonged to 24 families. After finishing her thesis, she left the clinic in 1906 (Wieser, 2001, p. 183).

Fürst reported her results in a paper published in 1907, in which she analyzed nine of her 24 families, the subjects ranging from 9 to 74 years of age. Important to our aims, she introduced many variations to Jung and Riklin's framework and technique. Although she intended to study normal individuals, Fürst quickly found that several family members evidenced very low intelligence, some even showing "imbecile traits," while others showed hysterical symptoms and the older ones signs of senility. These signs clearly pertained to the kind of eccentricities that Jung and Riklin identified with abnormality and should have led to those subjects being excluded from the sample. However, Fürst did not attempt to control variables such as intellectual level or mental functioning: given they "all were capable of living in society," she deemed the 100 individuals as a suitable sample. In brief, Fürst equated mental normality with successful social adaptation (Fürst, 1910/1918, pp. 408–413). She also altered Jung and Riklin's stipulation of a single standardized list of 400 words for an entire sample: she administered 400 words to 10 individuals and 200 to the remaining 90 individuals. When classifying the reactions to find group patterns Fürst adopted Jung and Riklin's scheme, but underlined the fact that classifying reactions did not shed light into the psychical processes behind word associations: classification "cannot of itself settle anything about the inner conditions of the association; it does not, indeed, deal with the question." Fürst chose to use the scheme anyways, not before underlining that she did so "because no better one [scheme] is known to me" (Fürst, 1910/1918, p. 409).

As for the experimental results, they were quite suggestive. At an individual level Fürst argued that most subjects of her sample reacted to a considerable part of the words in a constant manner. This meant that several individuals showed a preference for certain association forms—that is, inner, outer, sound associations, and so forth. Regarding group results, Fürst found that all children under 16 had more inner associations than their mothers, while all but one of the children over 16 had more outer associations that their mothers (Fürst, 1910/1918, p. 439). By using a mathematical formula devised by Jung, Fürst was able to conclude that there were strong resemblances among familiars in regards to their association styles or types.[13] For instance, in most cases where the mother tended to react through adjectives and value judgments, the daughter replicated the tendency.

As an explanation, Fürst first conjectured that "the female sex of the experimenter" could have "an effect upon the adjustment of the subject." Attention, Fürst argued, was but a fragment of affectivity, and sex differences could perhaps prompt different adjustments (i.e., reactions) in different subjects by influencing their attention— in this case, the attention of

[12] Fürst (Frl.) Emma, in the *Matrikeledition der Universität Zürich, 1833-1924*, Universitätsarchiv, Zurich. Retrieved from http://www.matrikel.uzh.ch/active/static/6822.htm. Fürst is completely absent from most historical accounts on psychoanalysis, even the ones focused on the life and work of early women psychoanalysts. See for instance Galatzer-Levy (2015) and Thompson (1987).

[13] The "average association difference" was defined as the average of the aggregated differences that existed for each kind of association between two given family members (Fürst, 1910/1918, p. 410).

mothers and daughters. It was conceivable that the subject "would have had a different attitude toward another experimenter." However, Fürst also conjectured that her results could be due to the distance between the experimenter's and the subject's intellectual and social "grade." Both alternatives suggested that Jung and Riklin's results had to be taken cautiously, given that they were both male and that many of their subjects were women who worked as their subordinates at the Burghölzli. In Fürst's terms, the subjects "were in the service of the experimenters; this influences the affectivity" (Fürst, 1910/1918, pp. 414–415, see also p. 442). Ultimately, she concluded that resemblances between mothers and daughters were due to "their common life and *milieu*" and the "mental dependence of the child" upon their mothers. The "relatively strong and uniform agreement" between mothers and children was seen by Fürst as her most important empirical finding.[14] However, her claims regarding the effect of interpersonal variables over associations represented an early critical comment on the objectivity and validity of Jung and Riklin's conclusions.[15]

In any case, the results of this first phase of experimental research at the Burghölzli pleased Bleuler. In 1904 he established a dedicated psychological laboratory at the Clinic at the back of the main building, near the institution's laundry and dairy. With Riklin gone, Jung was appointed director, and he continued the program as he saw fit (Clay, 2016, pp. 80–84; Kerr, 1994, p. 71).

RTs and the Reproduction Method (1903–1905)

Jung's experiments have been usually portrayed by historical scholarship as involving the measurement of the subject's RTs. However, between 1902 and 1903 WATs revolved exclusively on prompting associations and classifying the results. As stated by Fürst, her part of the experiments had begun in 1903, "when the importance of time in the reaction was not so well known" (Fürst, 1910/1918, p. 408). As an experimental feature, the measurement of reaction times was introduced by Jung and Riklin sometime later, probably between 1903 and 1904.[16] In fact, the topic was not publicly discussed by Jung until he wrote his 1905 *Habilitation* thesis on the issue (Jung, 1905/1918e; see also Loewenberg, 1995, p. 72).[17]

Jung's treatment of reaction times has been usually portrayed as a straightforward and simple matter of measuring seconds and gauging averages (e.g., Borch-Jacobsen & Shamdasani, 2013, pp. 60–61; Makari, 2008, pp. 256–268). The reality, however, was more complicated. Jung admitted that reacting to a given word was not a simple, two-step, stimulus-response phenomenon: following Édouard Claparède (1873–1940) and Theodor Ziehen, Jung acknowledged eight "components" in each psychological reaction, such as the transference of the sound to the subject's ear, the recognition and understanding of the word, the evocation of the induced presentation (the "pure association") and the excitation of the speech-motor apparatus. RTs were not simple, unitary magnitudes because their underlying processes were not simple, automatic phenomena. Jung chose to emphasize "some of the most important factors"—this is, both the evocation and the actual pronunciation of the reaction word—although he was clearly aware of the complexity of the topic (Jung, 1905/1918e, p. 227).

[14] Jung seems to have been positively impressed by Fürst's work, to the point he sent it to Freud. See Jung's letter to Freud on June 28, 1907 (McGuire, 1974, pp. 69–70).

[15] This criticism notwithstanding, Fürst's data greatly allowed Jung to advance his psychological theories about the social and familiar roots of the neuroses, for example at his Clark lectures (Jung, 1910).

[16] By May 1903, Jung was still not measuring RTs (Jung, 1906/1973b, pp. 166–172).

[17] Jung had begun working in his thesis in mid-1904; according to Bleuler, Jung was assisted by his then-patient Sabina Spielrein. See the letter from Bleuler to Mr. Spielrein on October 25, 1904 (Minder, 2001a, p. 32).

Jung measured reaction times in one-fifths of a second with a hand stop-watch—a rather gross measuring apparatus if compared with other, more sensible and "exact instruments" available at the time, such as the chronoscope (Jung, 1905/1973d, p. 329). According to Jung's philosophy of science, proving or disproving associations laws required a quantitative and accurate depiction of verbal reactions. However, subjects varied greatly in their reaction times, and while some of these variations involved several seconds, others involved smaller time fractions. Jung opted for using the stop-watch by acknowledging that the *clinical aim* of the experiments allowed "ignor[ing] slight differences so long as the causes of the greater differences do not escape us" (Jung, 1905/1918e, p. 228).

A third methodological choice involved statistical measurements. Jung measured the RTs of 26 of the original 38 subjects that had comprised the sample from his first study, collecting 4,144 individual magnitudes. How to make sense of these quantifications? As noted above, individual differences among the subjects and wide variations in the reaction times of a given individual greatly extended the quantitative range of the measurements. Isolated high values affected "the otherwise low average value in a very disturbing way" and made the arithmetical mean "quite misleading" (Jung, 1905/1918e, p. 230). Instead, he championed the use of the *probable* mean—the value that divided the RTs arranged in an increasing series into two halves. This value is what we would now call the "median."

Jung's statistical choices allowed him to identify the subject's actual pattern of reaction times in a more reliable fashion. The probable mean was a comparatively more representative figure of the *actual* psychical processes behind the mathematical figures. Besides reliability and validity, the probable mean was more *efficient* than other measures, such as Ziehen's "representative value": while these alternatives often required complex and obscure calculations, the probable mean could be gauged "more quickly," it only required mechanical skills (arranging the reaction times into a series) and could be applied to "great numbers of figures" (Jung, 1905/1918e, p. 230). This made them especially useful in clinical settings, where the physician using the test did not seek statistical rigor but differential diagnosis and getting acquainted with the subject's particular psychology.

Although interested in idiographics, Jung nevertheless still held a nomothetic ideal of science. Generalizations were needed to discover the universal laws guiding psychological associations. To combine both aims, Jung alternated between statistical measurements. Instead of pooling together the 4,144 measurements of his 26 subjects and calculating the entire series' average, Jung first calculated the probable mean—the median—for *each* subject. Only then did he calculate the arithmetical mean by pooling his 26 median values and obtaining an average. As a result, in 1905 Jung concluded that the universal mean duration of an association—the *temporal* aspect of the law of association—was of 1.8 s.[18]

Jung was skeptical of his own finding. He noted that it was a "pretty long duration," exceeding "by quite a good deal" the values given by previous literature.[19] He explained that several factors, such as the stop-watch's inaccuracy or the uneducated nature of the subjects of his sample that came from "low classes" could have delayed the reactions. Jung also hypothesized that idiomatic issues and language differences between Swiss and German subjects could be interfering. Linguistic factors clearly played a role: Jung found that the grammatical form of the stimulus word influenced both the form of the reactions and the RT itself.

[18] Jung's actual handling of statistical measurements renders untenable Hoffer's (2001, p. 119) opinion that his "early inclination was to go from the general to the individual." This also contradicts Cohen's (2014, p. 4) claim that the WAT produced "results upon which Jung declined to generalize."

[19] In 1879, Galton had informed that the average RT was 1.3 s, while in 1892 Féré had informed that average RTs was 0.7 for men and 0.83 for women (Jung, 1905/1918e, pp. 229–230).

Universal concepts required longer times to be associated with a word than concrete terms. Finally, Jung also found several unexplained overlong RTs.[20]

To make sense of the delayed reactions, Jung recurred to his contemporaries' ideas on emotionally toned complexes, as well as his own ideas on the matter.[21] He argued that reaction times were prolonged when conation (the affective or emotional aspect of mind) interpolated itself between the word and the reaction. If a stimulus aroused an emotionally accentuated complex, then the RT was considerably increased compared with the other associations.[22]

Jung's discoveries on reaction times had *clinical* implications: given overlong times denoted the presence of emotional complexes, then psychotherapists now had "the means by a short and simple examination to discover things of personal import, especially the complexes which characterize the psychology of the individual" (Jung, 1905/1918e, p. 239). Therapists could explore and diagnose pathology (e.g., in regards to hysteria), in which abnormally long RTs functioned as "valuable sign-posts" for "the discovery of the complexes making the disease, of which the hysteric is himself not always aware" (Jung, 1905/1918e, p. 239). In other words, it is clear that between 1904 and 1905 the word association *experiments* were transitioning into word association *tests.*

To check his ideas, Jung asked some of his original subjects for introspective reports of their associations. He also prompted further word associations, modifying the original word list on the fly, and explored the biographic and emotional background for overlong reaction times (Jung, 1905/1918e, pp. 240–259). Jung confirmed his suspicions: associations were slower when the word aroused an emotionally toned complex. However, the associations that directly *followed* the sensible reaction also had prolonged reaction times. As such, the emotional complex seemed to "bleed over" several associations as a "consequence of the maintenance of the emotional tone" (Jung, 1905/1918e, p. 246). This last discovery inadvertently exposed a weakness in Jung's outlook that would be exploited by critics such as Aschaffenburg (1906): if emotional tone dripped and influenced several reactions, and if a clinical exploration or examination was needed to find the critical reaction, then how could the physician identify the triggering word without suggesting anything to the subject or patient? In other words, there was no reliable way to identify which word had actually triggered each complex, making the identification of the complex itself a troublesome issue (Jung, 1905/1918e, p. 259).

In mid-1905 Jung devised what has retrospectively become one of the association experiments' most overlooked innovation: the reproduction experiment (Jung, 1905/1973a). According to Jung, some of his experimental subjects, especially those who could be diagnosed as "hysterics," often forgot certain stimulus words, or took long reaction times before uttering a word. He hypothesized that these phenomena could be complex-indicators pointing

[20] The standard to which one RT was compared with be called "overlong" was not the *group* average, but the individual subject's probable mean.

[21] Ziehen (1892) had demonstrated that the recollection of unpleasant memories required longer RTs than pleasant or neutral ones. Jung clearly benefited from this idea.

[22] It has been argued that Jung's test failed even at differentiating sexes (e.g., Borch-Jacobsen & Shamdasani, 2013, p. 60). However, by controlling further variables Jung did in fact draw conclusions from group statistics. Regarding the content of complexes, women had simpler complexes, most of them of erotic nature, while men were constellated by ambition and power. Moreover, gender influenced RTs: women reacted considerably more slowly to his chosen stimulus words. Finally, perseverations under distraction of attention increased in women and decreased in men (Jung, 1903/1973f, pp. 422–423; Jung, 1907/1973e, pp. 21–22).

to psychical repressions.[23] If such a hypothesis was true, then a correlation was to be expected between those associations that triggered complex indicators and later recollection failures. To check this idea, Jung added a further task to the WAT: the "new method" of reproduction (Jung, 1905/1973a, p. 279).

After having prompted the patient's associations and examining his or her complexes through "supplementary questions," Jung suggested the clinician could repeat the original word list as it had been presented to the subject (Jung, 1905/1973a, p. 279). The clinician then asked the subject to recall the original reactions, writing down the new reactions without measuring reaction times. Successes or failures could then be established through comparisons between the new reactions and the original ones. A subject succeeded if the original word was correctly recalled; he or she failed if his or her memory faltered, hesitated, or replaced the word (Jung, 1907/1918c). By way of this technique, Jung tested two patients and found that words that failed to be reproduced often, but not always, had followed overlong reaction times (Jung, 1906/1918d).[24] Hence, failures in the reproduction test had theoretical and clinical importance: they too pointed to repressed complexes. Incorrect reproductions were "directly constellated by a feeling-toned complex" or "immediately followed [by] a critical [reaction]" (Jung, 1905/1973a, pp. 278–279). As such, the reproduction experiment assisted in the closer circumscribing of complex disturbances by producing "untainted evidence" through an allegedly purely objective way (Jung, 1907/1918c, p. 396).

Thus, by 1905 Jung had proposed a relatively straightforward method for using the experiments as a test, while also producing evidence that suggested the test results were ambiguous and even contradictory. Contradictory evidence notwithstanding, by 1905 Jung concluded that there were 11 clear signs of an active constellating complex (or "complex indicators"). The most important ones were (1) a long RT, (2) a fault –the incapacity of coming up with a reaction word, (3) a perseveration, (4) a slip, and (5) a failure in the reproduction test. These indicators were the operationalization of the clinical phenomena that interested Jung and his colleagues in the context of experimental psychopathology. However, their apparent validity would soon come under attack from both inside and outside the Burghölzli.

Mental Adjustment Leads to Test Adjustment: Association in Abnormal Individuals (1904–1906)

The use of the association experiment as a clinical test was built over the conclusions drawn from the experiment's basic data. It is only natural that the test changed as the experimental results varied or were debated. The highly flexible and plastic nature of the test became evident as early as 1904, when a series of further findings forced Jung to reconsider the validity of the experiments as a whole.

As mentioned above, the experiments' original aim was to devise a reliable means for psychological differential diagnosis. Jung and Riklin had first focused on the associations of

[23] It has been claimed that Jung concluded that delays in reactions were due to repression in the Freudian sense (e.g., Clay, 2016, p. 86; Makari, 2008, p. 194). Other accounts have claimed that Jung was attempting to test Freud's hypotheses (Rosenzweig, 1992, pp. 15–17, p. 34). However, Jung clarified that, unlike Freud, by repression he meant "a more passive transference to the background." Indeed, Jung resorted to Bleuler's perspective to argue that it was "perfectly indifferent whether a mental process is conscious or unconscious" (Jung, 1905/1918e, p. 258).

[24] As evidence, Jung claimed that around 75% of the subjects' *incorrect* reproductions during the reproduction experiment concerned those reactions that were constellated by a complex (Jung, 1905/1973a, pp. 284–285; Jung, 1907/1918c, p. 405).

normal individuals to define a psychological standard. Shortly thereafter, the Burghölzli team focused on how the associations fared in *abnormal* clinical populations. A first inquiry was conducted by Kurt Wehrlin (1878–1966), another young assistant physician who worked at the Clinic from 1902 to 1904 and who carried his dissertation under Jung.[25]

For his doctoral research, Wehrlin collected the associations of 27 subjects from 17 to 69 years of age that pertained to the two lower groups of feeble-minded individuals: "imbeciles" and "idiots." However, the young physician found several methodological difficulties when testing feeble-minded individuals (Wehrlin, 1904–1905/1918, pp. 173–175). He had to administer the SWs in a slow manner, and during periods that ranged from several days to several weeks: changes which both influenced the results and precluded comparisons with other normal and abnormal samples. Moreover, given that certain words were clearly misunderstood by imbeciles, there was no way to guarantee that their association processes were operating with the intended stimuli: words "which are somewhat rare are therefore very badly understood and arouse much fewer related images because they are hardly known to the imbecile" (Wehrlin, 1904–1905/1918, pp. 178–179).[26] The subjects produced several "awkward and clumsy reactions" for each word and it was the experimenter who had to decide which one was to be written down in the experimental report; this introduced a strong subjective factor in the data gathering process. Finally, imbeciles were considered to suffer from acute degeneration of emotional functions, which modulated the complexes' emotional tone: as a consequence, any theoretical advance extracted from the data had limited generalizability.

These difficulties forced Wehrlin to alter Jung and Riklin's technique. He started out by claiming that his subjects' associations followed the law of contiguity or of similarity just as in normal subjects— strictly an a priori, unwarranted assumption. During the experiments Wehrlin constantly warned his subjects about potential misunderstandings, and he asked them "to give a reaction in one single 'appropriate' word." Such a procedural change, as he himself admitted, caused the experiment to "[lose] that freedom of its conditions by which what is specific can be brought out," further limiting the validity of the results (Wehrlin, 1904–1905/1918, p. 183). The assistant physician also modified Jung and Riklin's classification of reactions. Wehrlin found that their classification "was ill-adapted to the associations of imbeciles" and that "even where it could be used it did not bring out what was of pathognomonic importance" (Wehrlin, 1904–1905/1918, p. 178). As a result, Wehrlin *inductively* built a classification of associations drawing from his subjects' reactions. This classification was concededly simpler: it posed the existence of three broad imbecile reaction types. However, it had added clinical value precisely because it was more sensible to the real subject's actual disturbances and because it matched the clinical and symptomatologic classification. Wehrlin clearly put clinical utility over experimental rigor: "Our experiment is not designed to test the working capacity of the subject in any definite direction, but to create a method by which what is valuable for diagnosis may be brought out quickly and with certainty" (Wehrlin, 1904–1905/1918, pp. 183–184).

Jung himself was in charge of studying associations in epileptics, schizophrenics and neurotics at the Burghölzli. In 1905 "epileptic degeneracy" was defined by two features: intellectual symptoms such as feeble-mindedness and mental retardation, and emotional symptoms such as irritability, egocentricity, and exaggeration of feelings (Jung, 1905/1918a). As such, the association experiment was seen as an ideal technique to explore the epileptic's inner

[25] Wehrlin Kurt (Theodor), in the *Matrikeledition der Universität Zürich, 1833-1924*, Universitätsarchiv, Zurich. Retrieved from http://www.matrikel.uzh.ch/active/static/23457.htm.

[26] Distinguishing between delays, misunderstandings, and lack of reactions (faults) in feeble-minded associations was a major obstacle.

psychodynamics. However, epileptics were scarce at the Burghölzli, so to obtain a workable sample, Jung had to collaborate with the Swiss Asylum for Epileptics, depicted in Figure 1.

Jung and the superintendent of the Asylum gathered 18.277 associations from a total of 158 patients. However, as had happened to Wehrlin, Jung came across several methodological and procedural problems. Epileptics became "easily confused" when they were explained the experiment, and the emotional outbursts that followed "affect the result most distinctly" (Jung, 1905/1918a, p. 209). This problem was also found when testing schizophrenics: it was very difficult to decide whether they "simply cannot or will not answer," making difficult obtaining "the correct information" (Jung, 1907/1973e, p. 109).[27] These obstructions forced Jung to change the experimental task itself: he instructed the epileptics that they had to answer as quickly as possible "without thinking at all about it [the SW]." In some severe cases Jung even had to represent the experiment "as a kind of game of thoughts" (Jung, 1905/1918a, p. 210). Jung was aware of the fact that different requests prompted different kinds of mental phenomena in the experimental subject, so he knew that such changes undoubtedly altered the validity of the results.[28] He also had to put together a different word list, comprising 200 concrete and abstract terms, adjectives, and nouns.

Jung's results were suggestive. Epileptic associations resembled those made by the feeble-minded in several respects. For instance, both groups attempted to define the SW or to explain it through tautologies, and both types tended to answer in sentence form. In both groups the stimulus word sparked very intense reactions, prompting a large number of associations. Finally, both groups had abnormally long RTs, in some cases even reaching a 9-s delay. This suggested that the word association experiment *alone* could not distinguish between epileptics and feeble-minded: a deeper clinical examination was required. More striking, however, was the finding that epileptics also shared traits with normal people: neither normal nor epileptics tended to provide superficial word-associations, and in both groups the associations were in part "constellated by a disease complex"—meaning that not only epileptics had unconscious, dynamic complexes. Finally, those reactions that seemed typical of epileptics required a great deal of clinical interpretation and insight if they were to be of any practical use. For example, Jung noted that epileptics tended to explain their associations "in an awkward, circumlocutial [sic] character." However, both normal and schizophrenic individuals provided these explanations too (Jung, 1907/1973e, pp. 108–110). Besides, deciding which association qualified as awkward required making a clinical judgment based on an interpretation on the individual's inner life and personality. Jung concluded that he could "not venture to draw any universal deduction" from his epileptics, even admitting that his cases could be "exceptional," as in nonrepresentative (Jung, 1905/1918a, p. 226).

Jung was also interested in testing his association experiments in neurotic patients. He had begun collecting associations from hysterics since at least 1901 (Minder, 2001b, pp. 56–57).[29] He tested several neurotic patients between 1904 and 1905, publishing two papers on the issue in 1906. The first work was a broad presentation about the relationship between psychoanalysis and the association experiments as illustrated by a case of obsessional neurosis. In the paper, Jung compared his test with Freud's free association technique, arguing the WAT was more efficient and had less requirements than Freud's method (Jung, 1906/1918c).

[27] The case of dementia praecox reported in Jung's monograph involved a list of 82 SWs.

[28] In 1904 Jung had stated that "in our association experiment we tend to produce an exclusive excitation of the speech organism. The more exclusive this excitation is the more strongly will verbal reactions occur in the reaction" (Jung & Riklin, 1904/1918, p. 14).

[29] However, Jung began publishing results on hysterics only in 1903 (Jung, 1906/1973b). Incidentally, the 1903 article does not describe any list of 100 SWs, and reports a total of 324 associations.

According to Jung, hysterics were defined by the presence of a complex "endowed with extremely strong emotional tones, possessing such constellating power that it brings the whole individual under its influence," which made them ideal subjects for association testing (Jung, 1906/1918d, p. 299). This was also true for obsessive neurotics, who were "incapable of holding in their images with a tight rein" (Jung, 1906/1918d, p. 321). Thus, prompted associations clarified the nature of the complex and provided clues and suggestions for causal therapy.[30] Finally, the test also contributed to research by producing "scientific knowledge" on the origin and inner construction of the psychogenic neuroses (Jung, 1906/1918d, p. 300).

As an illustration of the interplay between the clinical and the experimental use of word association, Jung reported a case of obsessive neurosis he had treated in June 1905. Jung had subjected the female patient to a list of 100 words. The subject's probable mean RT was of 2.4 s. Given this figure was longer than the 1.8 s normal RT, Jung's first clinical conclusion was that the subject had "a marked emotivity" (Jung, 1906/1918d, p. 302). By tracking down perseverant reactions and partial amnesias Jung argued that the subject's ideas revolved around an erotic complex which involved self-reproach and self-criticism (Jung, 1906/1918d, pp. 310–321). Jung then confirmed the hypothesis by conducting a psychoanalysis through free association and dream analysis.

Jung provided further evidence for his claims in his second paper of 1906, which dealt on the relations between association, dreams and hysterical symptoms in a 24-year-old hysteric patient in the autumn of 1905 (Jung, 1906/1918b). As with his obsessive patient, Jung administered this young woman a list of 100 SWs. According to our psychiatrist, the predominance of outer associations and faults (74% of the total) was "striking," while the subject's probable mean RT, 5.2 s, suggested a strong distraction of attention during the test (Jung, 1906/1918b, p. 358). Jung initially conjectured this was due to absorption by a pathological complex. By prompting the subject to explain just what diluted her attention, Jung found that she contended herself "simply with apprehending the outer form of the words" because she was fixated on her physical symptoms (Jung, 1906/1918b, p. 359).

Jung tested the subject repeatedly over the weeks the analysis lasted, comparing the results obtained in each iteration. As the young woman's health improved, she produced more inner and less outer and sound associations, and her RTs shortened. This suggested a correlation between the physician's clinical appraisal and the test's objective results. Jung was also inclined to conclude that the WAT was sensible to personality *change*. However, there were other, more plausible explanations, such as the effect of practice: an effect first discovered by Wehrlin. In his 1904 paper Wehrlin (1904–1905/1918, pp. 201–202) had demonstrated that for a first series of stimulus words the average RT of his feeble-minded subjects was approximately five times longer than the normal average. However, if the experiment was repeated, the subjects reacted almost three times quicker than the first time—even if a different series of words was used! Thus, Jung's therapeutic results could be due to test habituation.

However, by 1906 Wehrlin was not the only source of criticism toward Jung's outlook and technique. Riklin himself had been studying and treating hysterics at the Burghölzli since 1901. In May 1904, just before he had left the Burghölzli for Rheinau, Riklin had read a paper at the Annual Meeting of the Association of Swiss Psychiatrists where he argued that the WAT "prove[d] that there is a type of hysterical reaction which can be distinguished from the normal and from other psychoses" (Riklin, 1904). He pursued this idea during his tenure at Rheinau, and reported his results in a paper published in 1906 where he dealt with eight institutional cases of hysteria.

[30] Jung granted that to detect complexes the clinician could "intersperse additional pertinent stimulus-words": this is, words that were not on any previous published list (Jung, 1905/1973a, p. 273; see also Jung, 1905/1973d, p. 334).

In a broad sense, Riklin's analysis supported Jung's: the WAT allowed tapping both sexual and general unconscious complexes which "r[a]n like a thread through all the reactions" (Riklin, 1918, p. 323). Riklin also used his and Jung's normal standards to highlight and diagnose abnormalities (see Riklin, 1918, p. 328, p. 352), so he clearly validated Jung's ideas to some extent.[31] More important to our aims, however, is the fact that Riklin subjected Jung's experimental approach to further methodological and procedural changes. As Wehrlin and Fürst had done, Riklin oscillated between administering 100 and 200 words, adapting the length to each case. In an even further deviation from the standard procedure, Riklin also changed the *language* of the words, for instance administering them in Russian when the patient was of Russian origin (Riklin, 1918, p. 323). He did not resort to Jung's reproduction method either, and claimed that measuring reaction times was unnecessary, for example when it came to his only male hysterical patient, a 23 year-old arsonist whose associations required not a quantitative measurement but a qualitative, clinical analysis (Riklin, 1918, pp. 347–349). This rather striking choice meant that Jung's most dear complex-indicator was neither indispensable nor mandatory for clinical examinations. Finally, Riklin suggested new complex-indicators, such as being distracted by the surroundings and suffering consciousness disturbances (Riklin, 1918, p. 353). Thus, already by 1906 the WAT was being heavily modified to suit a wide array of practical demands.

Methodological Triangulation, or the Last Nail in the Coffin: Word Associations and the Psycho-Galvanic Reflex, 1906–1909

In 1906 a further methodological innovation was introduced at the Burghölzli: the study of the psycho-galvanic reflex (hereafter, PGR). Jung first became aware of the PGR thanks to a suggestion made by Swiss neurologist Otto Veraguth (1870–1944), who had been using a specific apparatus, the galvanometer, to register changes in a subject's cutaneous electrical conductivity when exposed to emotional stimuli. The galvanometer measured the skin's electrical reaction with high precision: any change in the electric conductivity of the body caused a "deviation" or "excursion" of a beam of light projected over a scale. After Veraguth demonstrated the phenomenon to Jung, the later took interest in using the apparatus together with the WAT. As shown in Figure 2, Jung made slight modifications to the galvanometer and began "to experiment on his own account" (Jung & Peterson, 1973, p. 495).

The galvanometer further complicated Jung's research program, prompting even more debate about its coherence.[32] From 1906 to 1908 Jung put two of his young doctoral students, Ludwig Binswanger (1881–1966) and Hermann Nunberg (1884–1970), to conduct PGR-related research for their dissertations. A confessed admirer of Bleuler's doctrines, the 25-year-old Binswanger had joined the Clinic in June 1906 as an assistant physician (Binswanger, 1957, pp. 1–2). He carried out his doctoral research during the second half of 1906, and after finishing it in early 1907 he left the Burghölzli to work for his uncle, Otto Binswanger, in the Psychiatric University Clinic in Jena.[33]

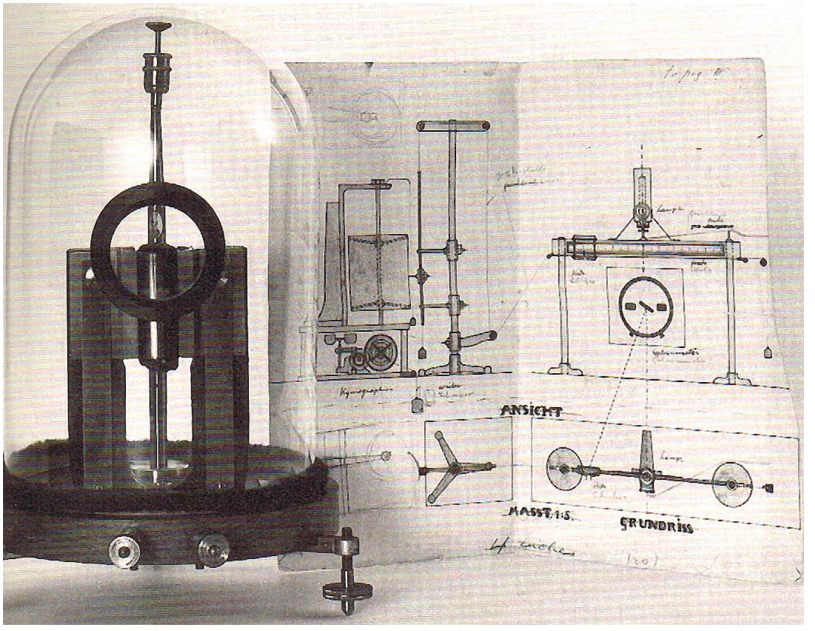
[31] In light of these findings, the widespread belief that holds that the WAT did not allow distinguishing between different pathological types or, even more poignantly, that it was a diagnostic "abject failure" (Borch-Jacobsen & Shamdasani, 2013, p. 60) should be nuanced.

[32] As stated, the galvanometer was introduced at the Clinic in 1906, not earlier. This must be noted to avoid conflating the different phases of Jung's research program, as it has been done before (e.g., Borch-Jacobsen & Shamdasani, 2013, p. 61; Clay, 2016, p. 13; Cohen, 2014, pp. 4–6).

[33] Binswanger stayed at the Jena Clinic between April 1907 and June 1908 (Falzeder, 2002, p. 87, n. 11; see also Jung's letter to Freud on 28 June, 1907 (McGuire, 1974, p. 68).

Figure 2

The Original Galvanometer Used by Carl Jung and His Collaborators at the Burghözli From 1906 Onwards



Note. The blueprint of the apparatus, which was devised by Jung himself, is in the background. *Source:* Gerhard Wehr (1989). C. G. Jung. René Coeckelberghs. Wikimedia Commons.

Ever since his arrival Binswanger championed the use of the WAT: he deemed it not as "an arbitrary means of expression," but as a method for exploring mental phenomena of theoretical and clinical relevance (Binswanger, 1907–1908/1918, p. 504). Consequently, he implemented the galvanometer as a methodological triangulation to clarify the affective processes involved in associations: if the PGR triggered by specific words could be measured, then both the emotional nature of reactions during the test as well as the scientific utility of the galvanometer could be demonstrated.[34]

To test his hypothesis, the young psychiatrist analyzed 2,160 reactions collected from 29 subjects, Jung himself among them, using statistical measurements (probable means) for analyzing reaction times and galvanic deviations.[35] However, Binswanger changed Jung's list of SWs on the fly to control the effect of the subjects' habituation. He then submitted his subjects to individualized psychological examinations to explore the origin and meaning of each association. Finally, to facilitate further analyses and comparisons Binswanger condensed each of the subjects' data on a single graphical representation. As shown in Figure 3, which coincidentally represents the data provided by Jung as an experimental subject, Binswanger

[34] Binswanger adopted Bleuler's theory of affectivity: the term "affect" encompassed not only affects "in the ordinary sense" but also "slight feelings or tones of pleasure and unpleasure [sic] at every possible kind of experience" (Binswanger, 1907–1908/1918, p. 455).

[35] Binswanger's subject for his first and fourth experiments was Jung himself (Binswanger, 1907–1908/1918, pp. 457–478, pp. 498–504). See also Figure 3 of this article.

Figure 3

Reaction Times (RTs) and Galvanic Deflections

Note. The marks placed on the horizontal line (A) represent the RTs for each stimulus word. The figures at the foot of the ordinate (B) represent RTs in one-fifth of a second. The bigger numbers at the ordinate axis (C) represent the galvanic deflections in centimeters and millimeters. The marks at the bottom of each of the bars that compose the rising or falling lines (D) represent the galvanometer deviations. *Source:* C. Jung (Ed.) (1910). Diagnostische Assoziationsstudien, p. 194. Verlag. Image in Public Domain.

perfected his director's own method of data presentation: the new "bar method" allowed for a more efficient and even intuitive clinical analysis of RTs, galvanic deviations and underlying complexes in any given individual (Binswanger, 1907–1908/1918, p. 459, see also p. 453).[36]

Binswanger's study confirmed some of Jung's ideas. Purely intellectual tasks such as adding or multiplying did not alter the skin's galvanization. Conversely, words that aroused emotionally toned complexes were accompanied by a change in bodily innervations and by higher galvanic deviations (Binswanger, 1907–1908/1918, pp. 504–508). As a result, overlong deviations were defined as valuable complex indicators. However, many strong deviations picked up by the galvanometer had followed *short* RTs (Binswanger, 1907–1908/1918, p. 509). This finding indicated that long RTs could be due to emotional arousal but also due to "intellectual difficulties," to "the influence of perseveration" or to "verbal factors" (Binswanger, 1907–1908/1918, pp. 529–530). Binswanger reached three broad conclusions. First, too-long RTs had a problematic diagnostic value. Second, what RTs actually measured was "the conflict between the [complex's] perseverating emotional tone and the new reaction [to the stimulus word]" (Binswanger, 1907–1908/1918, p. 529). Third, galvanic deviations were a more reliable indicator than RTs. In brief, every finding pointed to the superiority of the PGR as a measurement of emotions and affects.

Of course, Binswanger's criticism on the meaning of reaction times was not new: already in 1904 Wehrlin had criticized Jung's claims on reaction times by noting they were heavily influenced not only by psychological disturbances but also by practice (Jung, 1907/1973e, pp. 45–52; Jung, 1906/1918b, pp. 356–377).[37] What was indeed new was the *nature* of the evidence: Binswanger was criticizing RTs on objective, publicly verifiable and psychophysical

[36] Compare with Jung and Riklin (1918, pp. 169–172).

[37] In fact, Jung had built a significant part of his first lecture as Privatdozent in October 1905 over the diagnostic quality of overlong RTs (Jung, 1903/1973f, pp. 417–418).

grounds, and by using an apparatus that Jung himself endorsed. Moreover, Binswanger's critique on reaction times was part of a broader, mostly implicit but nonetheless effective attack on Jung's entire program. Indeed, the young assistant took issue with the entirety of Jung's complex-indicators: several, if not all of these indicators could be triggered by external, nonemotional factors. In 1906 Jung himself had admitted that in hysterics and schizophrenics the examination situation itself was "a complex-stimulant" (Jung, 1907/1973e, p. 101). Binswanger now provided empirical data showing that the subject's expectancy to the stimulus word, or even accidental behaviors such as the experimenter raising his eyes and looking to the stopwatch influenced the subject's disposition and reactions (Binswanger, 1907–1908/1918, pp. 453–454). In short, environmental factors and uncontrollable variables were meddling with Jung's allegedly purely empirical procedure.

In a broader sense, Binswanger emphasized the *clinical* method even further than Jung, by claiming that the subjects' "idiosyncrasies" revealed "many individual exceptions to conclusions that are correct for most cases" (Binswanger, 1907–1908/1918, p. 469). As such, universal generalizations such as the ones Jung had proposed, and even the normal universal standards, were untenable. In 1904, Jung and Riklin themselves had admitted that the associations obtained during an experiment made sense only as a function of the "individual peculiarity of the subject" (Jung & Riklin, 1904/1918, p. 167). What Binswanger did 3 years later was to exploit such acknowledgment. In reference to that very peculiarity, he argued that associations varied in accordance to the "psychological character of the subject" (Binswanger, 1907–1908/1918, p. 479). Reactions to SWs were subjective, and although they suggested the influence of a complex, the complex itself and its contents were highly personal and demanded a first-hand acquaintance with the subject's psychical reality. Binswanger broadened Jung and Riklin's idea about individual peculiarity until it encompassed both the tested subject *and* the tester himself. By arguing that each experimenter at the Burghölzli had a different personal equation, he claimed that the average values or reaction times obtained by one researcher could not be directly compared with the results of another (Binswanger, 1907–1908/1918, p. 457). In a sense, each researcher had to establish the averages and mean values of both his subjects' reactions and his or her own perceptual shortcomings. This precluded experimental comparisons, and meant a setback to Jung's aim of finding universal regularities for experimental psychopathology.

Binswanger concluded that the psychogalvanic phenomenon rendered a "real service" in psychological investigation, precisely because it enlarged the clinician's knowledge "by an objective complex sign" and provided "more direct information about the affective reactions *than other complex signs*" (Binswanger, 1907–1908/1918, p. 504 [emphasis mine], see also p. 528). In other words, just a year shy after Jung's introduction of the galvanometer at the Burghölzli his first doctoral student had used it to revise and discuss practically every one of his director's tenets.

After Binswanger finished his dissertation, Nunberg started his own psychophysical research under Jung (Nunberg, 1918). Another historically neglected Burghölzli assistant, this Polish trainee had spent 2 years studying medicine in Krakow before moving to Zurich in the autumn of 1906. After two more years of medical studies, in 1908 the 24-year-old became an *Unterassistenten* at the Burghölzli. As the topic for his doctoral dissertation, Nunberg chose the physical accompaniments of association processes: he published his results in 1909, and after obtaining his doctorate in 1910, he left Zurich after accepting a post at the Breitenau clinic in Schaffhausen (Wieser, 2001, p. 45).

As Jung and Binswanger before him, Nunberg's starting point was the assumption that emotions are accompanied by definite physical and bodily changes.[38] Because complexes

[38] For Jung's theory of emotional states and bodily changes, see Jung (1907/1973e, pp. 40–42).

associated themselves to both ideas and to the ego, then arousing those complexes should produce detectable bodily changes. To test his idea, Nunberg combined "the association test as elaborated by Jung" with an "objective method" that allowed him to study "all physical symptoms through which the feelings are outwardly displayed" (Nunberg, 1918, p. 532). The objective method was a combination of three experimental apparatus: a machine that detected bodily tremors, a pneumograph for recording chest movements during respiration, and the galvanometer. If the galvanometer really was an indicator of affective processes, then all physical manifestations of affectivity "should run parallel with the fluctuations of the galvanometer, at least to a certain extent" (Nunberg, 1918, p. 550). Consequently, he correlated the occurrence of three physical symptoms (tremors, breathing, and the PGR) and the triggering of emotional complexes through a list of 25 SWs in six subjects.

Nunberg's findings were twofold: first, hand tremors were stronger and breathing was more strained and superficial when complexes had been aroused. Second, these two physical manifestations strongly correlated with stronger galvanic deflections. This was interpreted as conceding even more validity to the PGR as an index of emotional states. However, Nunberg emphasized that he had found several individual differences during his experiments. For instance, one of his subjects had stronger hand tremors when they were presented an *indifferent* stimulus—this is, a noncritical stimulus word. Such "contradictory behavior" in the sample could be due to the individual's personality or to "an exception conditioned by the experimental method." In other words, breathing, tremors, emotions, and the PGR did not follow a clear pattern across individuals.

Nunberg concluded the contradiction could not be fully explained, citing "lack of material" as a further reason (Nunberg, 1918, p. 537). His conclusion was as paradoxical as it was representative of the state of the Clinic's research program toward 1908: relationships between psychological complexes and bodily changes "are universal relationships," Nunberg argued, but if one proceeded to the "individual figures of the subjects separately," then the relationships became rather different. "There are exceptions, so than individual deviations as well as their total average are in some test cases greater with unconscious than with conscious complexes" (Nunberg, 1918, p. 556).

Binswanger's and Nunberg's papers revealed Jung's failed attempts to find regularities in an enormous and very heterogeneous mass of clinical data. Further complications soon arose from Jung's international collaborations. In late 1906 Adolf Meyer, who had been positively impressed by Jung's work, suggested fellow physician and Columbia professor of psychiatry Frederick Peterson (1859–1938) to travel to the Burghölzli and collaborate with Jung (Leys, 1985).[39] Peterson worked at the Burghölzli for several weeks during late 1906 and early 1907, and published his results in *Brain* on July 1907 (Jung & Peterson, 1973).

Like Binswanger, Peterson was not interested in the WAT per se but as a means to secure methodological triangulation when researching the influence of emotional states over the PGR and respiration rates in normal and insane individuals. To explore these relations Peterson used a series of 19 different emotional and intellectual stimuli, which ranged from asking for the subject's age, to prompting word associations, up to inquiring about the physique of the subject's couple. Results were again mostly inconclusive: in normal individuals the relation between the galvanic deviations and the respiratory curves was irregular and inconstant, while insane individuals showcased even more haphazard reactions. The only regularity found suggested that in most subjects the most personal and emotionally laden stimuli triggered the most "lively emotions" and provoked the highest galvanic deflections (Jung & Peterson, 1973, p. 502).

[39] For a brief but informative analysis on Peterson's life and work, see Taylor (1998, pp. 102–107).

What about the WAT itself? Peterson wording suggests that he did not think word association could be used as a diagnostic technique because it was not useful at distinguishing between individuals:

> Two persons, of the same social class, one intelligent, the other unintelligent, even with differences in the character of their intellectual development, may still produce similar associations, because language itself has many general word connections which are familiar to all sorts of individuals belonging to the same circle of society. (Jung & Peterson, 1973, p. 525)

Ultimately, Peterson acknowledged that the galvanometer (not the WAT) qualified as "a measurer of the amount of emotional tone" and became "a new instrument of precision in psychological research (Jung & Peterson, 1973, p. 499).

A collaboration between Jung and Charles Ricksher (1879–1943) carried out from late 1906 to mid-1907 yielded similar uncertain results. An assistant physician at the Danvers Insane Hospital in Massachusetts, Ricksher was mostly interested in the PGR in its relation to respiration and to emotional psychical complexes. Tellingly, he did not even attempt to use the WAT: instead, he spoke four short sentences or words to his experimental subjects which ranged from "trite" and conventional up to "critical" and personal observations (Ricksher & Jung, 1907, p. 194). The PGR was stronger when the verbal stimuli were more "personal," and some subjects slowed their respiration after hearing those personal observations. However, the exceptions to the rule were not insignificant. Whether the respiration was slowed or quickened during the PGR "seem[ed] to depend on the individual," and while in some subjects the respiratory rate increased during an emotional reaction, in others the rate decreased (Ricksher & Jung, 1907, p. 196).

As contradictory evidence mounted, Jung started to waive his claims. In 1906 he acknowledged that "we have still not yet succeeded in finding a method of classification that is in principle entirely satisfactory" (Jung, 1903/1973f, p. 413). He also acknowledged that sometimes long reaction-times were due to causes unrelated to feeling-toned complexes (Jung, 1903/1973f, p. 418). The following year, he admitted that overlong RTs did not directly correlate with disturbances in reproduction: four out of each 10 failed reproductions were either below or near the average RT. He also granted that there were no clear, stable ways in which the complex-indicators presented themselves. The problem seemed to lie in the test's accuracy and its reliability (Jung, 1907/1918c, p. 398).

Jung also admitted that the associations from "very uneducated" people were often undistinguishable from those by "mentally defective subjects" (Jung, 1903/1973f, p. 416). Furthermore, his preferred statistical measure also came under scrutiny. It had now become evident that the probable mean was useful in large series of numbers, but that it became "too inexact" with small series because it was "affected by slight accidents" (Jung, 1903/1973f, p. 416). Jung was forced to reintroduce the use of the arithmetical mean into the experiments, and in early 1908 the research program was in a standstill. Our young psychiatrist became clearly frustrated, and in a letter to Freud on February 1908 he described his paper with Ricksher as written in an "idiotic style and with garbled results" (McGuire, 1974, p. 125).

While Binswanger and Nunberg were conducting their research, Jung's personal and professional relationship with Bleuler and the staff were rapidly deteriorating.[40] In October 1908,

[40] Freud to Abraham, 29 September, 1908, in Falzeder (2002, p. 168). For historical evidence of Jung and Bleuler's estrangement, see also Abraham to Freud, 16 July, 1908 (Falzeder, 2002, pp. 146–149), and Jung's increasingly frustrated and exasperated correspondence to Freud, for instance the letters of August 19, 1907; November 8, 1907; March 11, 1908; June 19, 1908; and August 21, 1908 (McGuire, 1974, p. 78, p. 97, p. 135, p. 157, p. 170). For an overview, see Möller et al. (2002).

Bleuler dismissed Jung from his post, and in March 1909 Jung effectively severed his ties with the Clinic (Clay, 2016, pp. 149–153). From 1909 onward he would focus on his independent practice and his participation in the then emerging international psychoanalytic movement. He would not publish any new data on the experiments, and the test itself would not be the center of any of his future endeavors. Ironically, by the time Jung lectured American psychologists and psychiatrists about the association method at Clark, he had already dislodged himself from his own research project.[41]

Conclusion: Test, "Test," and Tests

A detailed historical analysis of the Zurich word-association experiments during the first decade of the 20th century reveals several significant points. Regarding their inception, it is clear that Jung worked alongside half a dozen colleagues in the development and refinement of association experiments. He was brought to the Burghölzli by Bleuler, who also paired him with Riklin. It was Riklin who both taught Jung the experimental procedure and carried out a seminal part of the experiments. And it was Jung's doctoral students—Wehrlin, Fürst, Binswanger, and Nunberg—who contributed both empirical data and criticism. As a result, Jung's program was a highly collaborative research effort.

Regarding the experiments themselves, they developed naturally into a psychological test: standard figures obtained from normal subjects were used in clinical populations for assessment and diagnostic purposes. By 1906, Jung clearly distinguished between collecting associations as a theory-producing device and prompting associations for diagnostic ends. But as the research program developed the Zurichers reached lesser and lesser consensus on the experiment's data, assumptions, and procedures. The classificatory scheme of associations, the experimental task and even the criteria for defining normalcy were debated. Results were inconclusive: associations were seen as influenced by the experimenter's sex, by the family group, by the social milieu, by psychopathological mental functioning, by the effect of practice, and even by accidental environmental stimuli. Neither the experimental meaning nor the clinical value of overlong reaction times could be settled, and by the time Jung left the Clinic his colleagues were emphasizing individual differences to the point they discouraged universal generalizations.

These disagreements precluded the WAT achieving a single, stable character. Several modifications were introduced by different psychiatrists working in disparate settings and pursuing diverse aims, leading to the test's high plasticity and low standardization. The amount and content of the SWs, the classification scheme, the reproduction method, the measurement of RTs, and even the order given by the experimenter were constantly modified, sometimes even by Jung himself. The only two constants were the test's general aim (exploring complexes through psychological examinations) and its fundamental task (uttering a word after hearing a stimulus).

Ultimately, it is clear that Swiss psychiatrists did not offer conclusive evidence on the test's scientific and professional value. However, their experiences in word association stimulated scholarly debate and cemented the Burghölzli's international reputation. Zurich psychiatrists opened many doors for North American psychologists by providing them with a much-needed scientific tool for conducting psychological examinations. The process by which these clinical psychologists circumvented the manifest limitations of the "test" and developed it in

[41] Loewenberg (1995, pp. 74–76). See also Kerr (1994, p. 280ff). Jung did keep developing his ideas on psychical complexes, but he did so on clinical and anthropological, not experimental, grounds; in none of his post-1909 publications did he present new data as obtained through the test.

novel directions will be the subject of a future study. For now, however, it should be noted that nothing prevents the word "failure" from being associated with "success."

References

Aschaffenburg, G. (1906). Die Beziehungen des sexuellen Lebens zur Entstehung von Nerven-und Geisteskrankheiten [The relations of the sexual life to the development of nervous and mental diseases]. *Munchener Medizinische Wochenschrift*, *53*, 1793–1798.

Baker, D. B., & Benjamin, L. T., Jr. (2014). *From seance to science: A history of the profession of psychology in America*. University of Akron Press.

Bernet, B. (2013). *Schizophrenie: Entstehung und Entwicklung eines psychiatrischen Krankheitsbildes um 1900* [Schizophrenia: Emergence and development of a psychiatric clinical picture around 1900]. Chronos Verlag.

Binswanger, L. (1918). The psychogalvanic phenomenon in association experiments. In C. G. Jung (Ed.), *Studies in word-association* (pp. 446–530). Moffat, Yard & Co. (Original work published 1907–1908)

Binswanger, L. (1957). *Sigmund Freud: Reminiscences of a friendship*. Grune and Stratton.

Bleuler, E. (1918a). Consciousness and association. In C. G. Jung (Ed.), *Studies in word-association* (pp. 266–296). Moffat, Yard & Co. (Original work published 1905)

Bleuler, E. (1918b). Upon the significance of association experiments. In C. G. Jung (Ed.), *Studies in word-association* (pp. 1–7). Moffat, Yard & Co. (Original work published 1904)

Borch-Jacobsen, M., & Shamdasani, S. (2013). *The Freud files*. Cambridge University Press.

Clay, C. (2016). *Labyrinths: Emma Jung, her marriage to Carl and the early years of psychoanalysis*. William Collins.

Cohen, E. D. (2014). *C. G. Jung and the scientific attitude*. Rowman & Littlefield.

Davidson, I. (2020). *Personal politics: The rise of personality traits in the century of eugenics and psychoanalysis* [Unpublished Doctoral Dissertation]. York University.

Ellenberger, H. (1994). *The discovery of the unconscious. The history and evolution of dynamic psychiatry*. Fontana Press. (Original work published 1970)

Evans, R. B., & Koelsch, W. A. (1985). Psychoanalysis arrives in America. The 1909 psychology conference at Clark University. *American Psychologist*, *40*(8), 942–948. https://doi.org/10.1037/0003-066X.40.8.942

Falzeder, E. (2015). Profession—psychoanalyst. In E. Falzeder (Ed.), *Psychoanalytic filiations: Mapping the psychoanalytic movement* (pp. 103–127). Karnac.

Falzeder, E. (Ed.). (2002). *The complete correspondence of Sigmund Freud and Karl Abraham: 1907-1925*. Karnac.

Frey-Rohn, L. (1990). *From Freud to Jung: Comparative study of the psychology of the unconscious*. Shambala.

Fürst, E. (1918). Statistical investigations on word associations and on familial agreement in reaction type among uneducated persons. In C. G. Jung (Ed.), *Studies in word-association* (pp. 407–445). Moffat, Yard & Co. (Original work published 1910)

Galatzer-Levy, R. M. (2015). Women and children last: Reflections on the history of child psychoanalysis. *The Psychoanalytic Study of the Child*, *69*, 108–145. https://doi.org/10.1080/00797308.2016.11785525

Graf-Nold, A. (2001). The Zürich School of Psychiatry in theory and practice. Sabina Spielrein's treatment at the Burghölzli Clinic in Zürich. *The Journal of Analytical Psychology*, *46*(1), 73–104. https://doi.org/10.1111/1465-5922.00216

Green, C. (2019). *Psychology and its cities: A new history of early American psychology*. Routledge.

Green, C. D., & Feinerer, I. (2016). The evolution of the American Journal of Psychology, 1904-1918: A network investigation. *The American Journal of Psychology*, *129*(2), 185–196. https://doi.org/10.5406/amerjpsyc.129.2.0185

Hoffer, A. (2001). Jung's analysis of Sabina Spielrein and his use of Freud's free association method. *The Journal of Analytical Psychology*, *46*(1), 117–128. https://doi.org/10.1111/1465-5922.00218

Homans, P. (1979). *Jung in context: Modernity and the making of a psychology*. University of Chicago Press.

Jehle-Wildberger, M. (2020). Adolf Keller's and C. G. Jung's development up to 1909. In M. Jehle-Wildberger (Ed.), *On theology and psychology. The correspondence of C. G. Jung and Adolf Keller* (pp. 3–15). Princeton University Press. https://doi.org/10.2307/j.ctvvb7mdd.6

Jung, C. G. (1910). The familiar constellations. *The American Journal of Psychology*, *21*(2), 240–251.

Jung, C. G. (1918a). Analysis of the associations of an epileptic. In C. G. Jung (Ed.), *Studies in word-association* (pp. 206–226). Moffat, Yard & Co. (Original work published in 1905)

Jung, C. G. (1918b). Association, dream and hysterical symptoms. In C. G. Jung (Ed.), *Studies in word-association* (pp. 354–395). Moffat, Yard & Co. (Original work published in 1906)

Jung, C. G. (1918c). On disturbances in reproduction in association experiments. In C. G. Jung (Ed.), *Studies in word-association* (pp. 396–406). Moffat, Yard & Co. (Original work published in 1907)

Jung, C. G. (1918d). Psycho-analysis and association experiments. In C. G. Jung (Ed.), *Studies in word-association* (pp. 297–321). Moffat, Yard & Co. (Original work published in 1906)

Jung, C. G. (1918e). Reaction time in association experiments. In C. G. Jung (Ed.), *Studies in word-association* (pp. 227–265). Moffat, Yard & Co. (Original work published in 1905)

Jung, C. G. (1973a). Experimental observations on the faculty of memory. In G. Adler & F. C. Hull (Eds.), *The collected works of C. G. Jung* (Vol. 2, pp. 272–287). Princeton University Press. (Original work published 1905)

Jung, C. G. (1973b). On simulated insanity. In G. Adler & F. C. Hull (Eds.), *The collected works of C. G. Jung* (Vol. 1, pp. 159–187). Princeton University Press. (Original work published in 1903)

Jung, C. G. (1973c). On the doctrine of complexes. In G. Adler & F. C. Hull (Eds.), *The collected works of C. G. Jung* (Vol. 2, pp. 598–604). Princeton University Press. (Original work published 1911)

Jung, C. G. (1973d). The psychological diagnosis of evidence. In G. Adler & F. C. Hull (Eds.), *The collected works of C. G. Jung* (Vol. 2, pp. 318–352). Princeton University Press. (Original work published 1905)

Jung, C. G., (1973e). The psychology of dementia praecox. In H. Read, M. Fordham, G. Adler, & W. McGuire (Eds.), *The collected works of C.G. Jung* (Vol. 8, pp. 1–151). Princeton University Press. (Original work published in 1907)

Jung, C. G. (1973f). The psychopathological significance of the association experiment. In G. Adler & F. C. Hull (Eds.), *The collected works of C. G. Jung* (Vol. 2, pp. 408–425). Princeton University Press. (Original work published 1906)

Jung, C. G., & Riklin, F. (1918). The association of normal subjects. In C. G. Jung (Ed.), *Studies in word-association* (pp. 8–172). Moffat, Yard & Co. (Original work published 1904)

Jung, C., & Peterson, F. (1973). Psychophysical investigations with the galvanometer and pneumograph in normal and insane individuals. In G. Adler & F. C. Hull (Eds.), *The collected works of C. G. Jung* (Vol. 2, pp. 492–553). Princeton University Press. (Original work published 1907)

Kerr, J. (1994). *A most dangerous method: the story of Jung, Freud, and Sabina Spielrein*. Knopf Doubleday.

Kirsch, T. (2004). History of analytical psychology. In J. Cambray & L. Carter (Eds.), *Analytical psychology. Contemporary perspectives in Jungian analysis* (pp. 5–31). Brunner-Routledge.

Kohs, S. C. (1914). The association method in its relation to the complex and complex indicators. *The American Journal of Psychology*, *25*(4), 544–594. https://doi.org/10.2307/1413291

Leys, R. (1981). Meyer's dealings with Jones: A chapter in the history of the American response to psychoanalysis. *Journal of the History of the Behavioral Sciences*, *17*(4), 445–465. https://doi.org/10.1002/1520-6696(198110)17:4<445::AID-JHBS2300170402>3.0.CO;2-9

Leys, R. (1985). Meyer, Jung, and the limits of association. *Bulletin of the History of Medicine*, *59*, 345–360.

Loewenberg, P. (1995). The creation of a scientific community: The Burghölzli, 1902-1914. In P. Lowenberg (Ed.), *Fantasy and reality in history* (pp. 46–89). Oxford University Press.

Makari, G. (2008). *Revolution in mind. The creation of psychoanalysis*. Harper Collins.

McGuire, W. (Ed.). (1974). *The Freud/Jung letters*. Princeton University Press.

Meyer, A. (1905). Review of Diagnostiscke Assoziationstudien. I. Beitrag. Experimentelle Untersuchungen über Assoziationen Gesunder. *Psychological Bulletin*, *2*(7), 242–250. https://doi.org/10.1037/h0065862

Meyer, A. (1906). Application of association studies. *Psychological Bulletin*, *3*(8), 275–280. https://doi.org/10.1037/h0068714

Meyer, A. (2013). *Sites of the unconscious: Hypnosis and the emergence of the psychoanalytic setting*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226058009.001.0001

Minder, B. (2001a). Burghölzli hospital records of Sabina Spielrein. *The Journal of Analytical Psychology*, *46*(1), 15–42. https://doi.org/10.1111/1465-5922.00213

Minder, B. (2001b). Sabina Spielrein. Jung's patient at the Burghölzli. *The Journal of Analytical Psychology*, *46*(1), 43–66. https://doi.org/10.1111/1465-5922.00214

Möller, A., Scharfetter, C., & Hell, D. (2002). Development and termination of the working relationship of C. G. Jung and Eugen Bleuler 1900-1909. *History of Psychiatry*, *13*(52), 445–453. https://doi.org/10.1177/0957154X0201305206

Monahan, P. A. (2009). C. G. Jung: Freud's heir or Janet's? The influence upon Jung of Janet's dissociationism. *International Journal of Jungian Studies*, *1*(1), 33–49. https://doi.org/10.1080/19409050802681876

Nunberg, H. (1918). On the physical accompaniments of association processes. In C. G. Jung (Ed.), *Studies in word-association* (pp. 531–560). Moffat, Yard & Co. (Original work published 1909)

Reisman, J. (1991). *A history of clinical psychology*. Taylor & Francis.

Ricksher, C., & Jung, C. G. (1907). Further investigations on the galvanic phenomenon and respiration in normal and insane individuals. *The Journal of Abnormal Psychology*, *2*(5), 189–217. https://doi.org/10.1037/h0073786

Riklin, F. (1904). Die diagnostische Bedeutung der Associationen bei der Hysterie [The diagnostic significance of associations in hysteria]. *Psychiatrisch-Neurologische Wochenschrift*, *29*, 275.

Riklin, F. (1918). Cases illustrating the phenomena of association in hysteria. In C. G. Jung (Ed.), *Studies in word-association* (pp. 322–353). Moffat, Yard & Co. (Original work published 1906)

Rosenzweig, S. (1992). *The historic expedition to America (1909). Freud, Jung and Hall the King-Maker*. Rana House.

Shamdasani, S. (1998). From Geneva to Zurich: Jung and French Switzerland. *The Journal of Analytical Psychology*, *43*(1), 115–126. https://doi.org/10.1111/1465-5922.00012

Shamdasani, S. (2004). *Jung and the making of modern psychology*. Cambridge University Press.

Shamdasani, S. (2011). Introduction: Jung, New York, 1912. In C. G. Jung & S. Shamdasani (Eds.), *Jung contra Freud: The 1912 New York lectures on the theory of psychoanalysis* (pp. vii–xxi). Princeton University Press.

Skues, R. (2012). Clark revisited: Reappraising Freud in America. In J. Burnham (Ed.), *After Freud left: A century of psychoanalysis in America* (pp. 49–84). Chicago University Press. https://doi.org/10.7208/chicago/9780226081397.003.0004

Smith, R. (2013). *Between nature and nurture. A history of psychology*. Reaktion.

Taylor, E. (1996). The new Jung scholarship. *Psychoanalytic Review*, *83*(4), 547–568.

Taylor, E. (1998). Jung before Freud, not Freud before Jung: The reception of Jung's work in American psychoanalytic circles between 1904 and 1909. *The Journal of Analytical Psychology*, *43*(1), 97–114. https://doi.org/10.1111/1465-5922.00011

Thompson, N. (1987). Early women psychoanalysts. *The International Review of Psycho-Analysis*, *14*, 391–407.

Walsh, R., Teo, T., & Baydala, A. (2014). *A critical history and philosophy of psychology. Diversity of context, thought, and practice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139046831

Wehrlin, K. (1918). The associations of imbeciles and idiots. In C. G. Jung (Ed.), *Studies in word-association* (pp. 173–205). Moffat, Yard & Co. (Original work published 1904–1905)

Wieser, A. (2001). *Zur frühen Psychoanalyse in Zürich: 1900–1914* [On Early Psychoanalysis in Zurich: 1900–1914] [Unpublished doctoral dissertation]. University of Zürich.

Zaretsky, E. (2004). *Secrets of the soul: A social and cultural history of psychoanalysis*. Knopf.

Zenderland, L. (2001). *Measuring minds: Henry Herbert Goddard and the origins of American intelligence testing*. Cambridge University Press.

Ziehen, T. (1892). *Introduction to physiological psychology*. Swan Sonnenschein & Co. https://doi.org/10.1037/12959-000ss

Received November 6, 2021
Revision received February 7, 2022
Accepted February 9, 2022 ■