Lang, J.W.B., & Corstjens, J. (in press). Key episodes in the history of testing in Central Western Europe. In S. Laher (Ed.), Comparative histories of psychological assessment. Cambridge, UK: Cambridge University Press.

Key Episodes in The History of Testing in Central Western Europe

Jonas W. B. Lang
Department of Human Resource Management and Organizational Psychology, Ghent University and Business School, University of Exeter

Jan Corstjens
Department of Human Resource Management and Organizational Psychology, Ghent University

This chapter focuses on the development of modern testing in the Dutch, German, and French-speaking parts of Europe (France, The Netherlands, Belgium, Luxembourg, Germany, most parts of Switzerland, Austria, and the South Tirol region of Italy). Some early forms of selection testing in this region of the world can be traced back as far as the Roman empire where officers and soldiers were selected through systematic assessment procedures during a period of probation before they were fully accepted into the Roman army (Stout, 2018). However, the start of modern psychometric testing happened much more recently toward the end of the 19th and early 20th century, and the region has played a pivotal role. The purpose of this chapter is to highlight some key moments in this history.

Wilhelm Wundt, Experimental Psychological Laboratories, and Early Testing

An important antecedent of the development of psychometric testing was the development of experimental psychology. In 1879, Wilhelm Wundt had opened the first psychological laboratory at the University of Leipzig, then a part of the newly formed German Empire. Wundt had worked with the famous psychophysiologist Hermann von Helmholtz at the University of Heidelberg and was interested and was interested in applying rigorous scientific measurement methods developed for psychophysiological measurement to the systematic measurement of psychological experiences. Soon after Wundt opened his laboratory in Leipzig, other psychological laboratories followed like the ones of Gustav Elias Müller in Göttingen and Hermann Ebbinghaus at the University of Berlin (Sprung & Sprung, 2001). Psychological laboratories typically had a set of devices to record reaction times, movements, or perceptions. One characteristic of early experimental psychology that resulted from its development from psychophysiology and psychology was a keen interest in the scientific measurement of psychological states. Early experimental psychologists defined psychological experiments by rigorous and repeatable measurement of psychological states and not necessarily by experimental manipulations (Stern, 1911; Wundt, 1883).

The development of experimental psychology was not restricted to its origin in Germany and soon spread to other countries like the United States and France. Frequently, foreign students traveled to Leipzig to study with Wundt and learn the new science of psychology. These students then took their knowledge and expertise home and psychological laboratories opened at universities like Clark (Nicolas & Young, 2014), Cornell, and later Harvard in the United States. Several famous early American and UK psychologists wrote their doctoral thesis with Wundt in

German. Notable figures include Charles Spearman from the UK and Titchener from the US. The founding father of psychology in the United States—William James—also studied briefly with Wundt but never worked in his laboratory (Gundlach, 2018). He later hired Hugo Münsterberg—also a former doctoral student of Wundt to built the psychological laboratory at Harvard University (Blatter, 2014). The new discipline was also well received in France, Belgium, and in the Netherlands where a psychological laboratories were opened at the Sorbonne University and the Salpêtrière University Hospital in Paris (Nicolas, 2006; Nicolas & Young, 2014), at the Catholic University Leuven in 1891 (Richelle, Janssen, & Bredart, 1992), and at the University of Groningen in 1892 (Busato, van Essen, & Koops, 2013).

From an early point in time, experimental psychologists that were trained and inspired by Wundt were interested in studying individual differences in mental capacity and talent. This work extended an article on the measurement of psychological states by Wundt himself (Wundt, 1883). The most well-known of these students was James McKeen Cattell. Cattell wrote a dissertation with the title *Psychometrische Studien* (Cattell, 1886a) that was also published in Wundt's journal *Philosophische Studien* (Cattell, 1886b). This early work studied relatively basic tasks, like for instance, the time it takes to recognize German words. In addition to the work of his doctoral advisor Wundt, Cattell also was inspired by the work of the English statistician and eugenicist Francis Galton who advocated the use of statistical methods and argued that human characteristics follow a bell-shaped normal distribution. Cattell later developed a series of basic cognitive measures at Columbia University in a first attempt to measure intelligence. However, one of his own students—Clark Wisseler—found evidence that these basic measures did not correlate highly with each other or with relevant outcome variables (Wissler, 1901) which was discouraging to Cattell and many of his contemporaries especially in the eugenics movement (Fancher, 1985; Wickett, 1998; Wissler, 1901).

Another Wundt-trained psychologist who contributed to the early development of psychological measurement was Charles Spearman. Spearman's seminal work on general intelligence (Spearman, 1904) was related to Cattell's earlier work and was also inspired by Galton. Both Cattell and Spearman relied on combining Galton's work with ideas around psychometric and psychological measurement that they had picked up in Wundt's lab in order to apply these insights in practice. Their focus on application differed from the perspective of their mentor who had high methodological and philosophical standards and believed that this type of application could only happen after the young discipline had developed—these views typically led to some distance between Wundt and the scholars he had trained (Fahrenberg, 2012). Spearman also focused on developing new statistical and psychometric methods like correlation coefficients and factor analysis methodology. He presented some of this work at the first congress of the newly founded German association for experimental psychology (later the German Psychology Association) in 1904 (Schumann, 2004).

Alfred Binet and the Measurement of Intelligence

In the 1891, the experimental psychological laboratory of the Sorbonne University in Paris was joined by an ambitious young researcher—Alfred Binet (see Figure 1). The Sorbonne laboratory had been founded and was led by the physiologist Henry Beaunis on the initiative of the founding father of French psychology Théodule Ribot. Binet had previously worked with Jean Martin Charcot at the Salpêtrière hospital. Charcot was famous for his work on hypnosis but he also had an interest in experimental psychology and had started his own psychological laboratory at the Salpêtrière university hospital. Binet quickly became an important member of the Sorbonne laboratory and finally its director in 1893. Binet was ambitious in trying to ensure

that French experimental psychology would catch up with German and American experimental psychology (Binet, 1894; Nicolas & Young, 2014). In addition to his work on hypnosis with Charcot, he had developed an interest in developmental psychology and had monitored the development of his two daughters (Siegler, 1992). However, the young Binet was not really satisfied with these earlier approaches and was thus interested in the new discipline of experimental psychology and the work in laboratories. In his laboratory work, Binet showed an interest in measurement early on and developed or refined several devices to measure basic individual differences (Nicolas, 2006). Binet was a member of the Society for the Psychological Study of the Child and in this capacity he and other members were approached by governmental Commission for the Retarded with the request to advise on a test for identifying children with learning disabilities. In response to this request, Binet collaborated with the then young medical student Théodore Simon and combined his earlier experiences from developmental psychology with his expertise with measuring psychological states. The result was the first modern approach to intelligence testing (Binet & Simon, 1907; Schneider, 1992). Binet and Simon's contribution was especially noteworthy because they went beyond earlier approaches that mainly focused on measuring perceptual speed and basic cognitive function. In contrast, Binet and Simon included more complex cognitive tasks into their scale. Binet and Simon were mostly concerned with individual diagnosis but in allowing for individual diagnosis they also gathered normative information for populations in the way Galton and his followers had recommended. Binet and Simon's work is remarkable because modern intelligence testing is still heavily based on the core task that they suggested more than a century ago. The Binet-Simon test became popular internationally because it solved the issues that Cattell was confronted with and thus was enthusiastically received and popularized especially in the US by Goddard and Terman (Benjamin, 2009; Terman, 1916). Although popular abroad, Binet's work gained less traction in France (Schneider, 1992). One potential reason was that French researchers and the French public in large parts were interested and fascinated by graphology (the idea that the style of handwriting – not its content – allows for insights into the writer's personality or psychological state). Binet initially rejected graphology but subsequently worked on graphology in collaboration with a graphologist in the years in which he also conducted work on intelligence in children (Nicolas, Andrieu, Sanitioso, Vincent, & Murray, 2015). Binet suggested the use of the Binet-Simon intelligence scale or similar measures in the military after reading about the use of tests with mentally ill German soldiers but those ideas gained not the same traction as in the US where a modified form of Binet's and Simon's test was used to systematically screen soldiers for world war I. One reason why intelligence testing was less popular in France in Binet's time and was largely used for children and psychiatric patients was the fact that France traditionally had a centralized system with a governmentally organized education system. This system included state exams. Because these exam scores were available for all recruits/school graduates and widely accepted there was less need to introduce a standardized measure in comparison to the US with its more decentralized and heterogenous educational system (Schneider, 1992).

William Stern, and the Foundation of Differential Psychology

A German psychologist who was inspired by Cattell's work on the application of statistical methods to testing was William Stern (see Figure 2). Stern had studied under Herman Ebbinghaus at one of the early experimental psychology laboratories at the University of Berlin and also had a keen interest in developmental psychology. Today, Stern is mostly known for suggesting the term "IQ" or "Intelligenzquotient" to track children's cognitive development (Stern, 1914). Arguably, his most important influence on the field of individual difference

research and testing is his work on differential psychology (Stern, 1900, 1911) and especially Stern's 1911 book "Die differentielle Psychologie in ihren methodischen Grundlagen". The book introduced key conceptual ideas on individual differences. While earlier work had typically discussed "measurement", Stern developed the conceptual foundation of modern individual difference and personality research. He first discussed traits ("Merkmale") and described differential psychology and testing as a separate applied discipline (see Figure 3). The book is less well-known outside the German speaking world because it was not translated but early individual difference researchers in other parts of the world read and cited it heavily. For instance, Gordon Allport was a student of William Stern in the 1920s in Hamburg and the first chapter of Gordon Allport's classic personality introduction book not only discusses Wundt's work but also many of Stern's ideas and cites both (Allport, 1937; Asendorpf, 1999).

Hugo Münsterberg and the First Selection Tests

Another former doctoral student of Wundt that was already briefly mentioned because of his association to Harvard and William James is Hugo Münsterberg (Benjamin, 2006; Blatter, 2014; Stern, 1917). Münsterberg wrote several influential books on the application of psychological methods. His 1912 book (Münsterberg, 1912) "Psychologie und Wirtschaftsleben: Ein Beitrag zur angewandten Experimental-Psychologie" is commonly seen as the starting point for modern industrial and organizational psychology and Münsterberg is therefore also typically seen as the or one of the two founding fathers of industrial and organizational psychology in the US (Spillmann & Spillmann, 1993; Vinchur & Koppes, 2007). The book was initially published in German and then translated by the author and published in English a year later as "Psychology and industrial efficiency (Münsterberg, 1913). Relevant to the history of testing, the 1912 book includes arguably one of the first successful applications of psychometric testing in an applied setting beyond educational and developmental contexts. A common problem at the time was the high number of traffic accidents. Modern cities increasingly adopted electric trams and these trams were responsible for a large number of traffic victims. Münsterberg successfully developed psychometric tests designed to select tram drivers and normed them on the basis of comparing tram drivers with many accidents with those with less accidents. Münsterberg also reported selection procedures for sailors and ideas about vocational tests. A new element in the 1912 book that went beyond earlier testing was the idea to match tests to the specific requirements of the job.

Otto Lipmann, William Stern, Institutes of Applied Psychology and the First Journal of Applied Psychology

While testing and particularly intelligence testing were less commonly used in the military, especially German psychologists initially enthusiastically applied testing in the military for more specific selection tasks, in organizations, and in vocational counseling (Gundlach, 1996, 2002; Vinchur & Koppes, 2010). William Stern is today widely known as an individual difference researcher and developmental psychologist. However, Stern was also an early industrial and organizational psychologist. Stern had come up with the German term "Psychotechnik" (Psychotechnology) that Münsterberg popularized in Germany (Gundlach, 1996, 2002) and captured the idea that psychology could be utilized as a form of technology. Stern extensively collaborated with Otto Lipmann—also a former student of Ebbinghaus. Lipmann and Stern led a trend toward the formation of a series of institutes of applied psychology in Germany. Most of these institutes sough collaboration with business organization but also more broadly with educational institutions and the government. Some of these institutes were linked to universities and directly sought to apply psychological methods and particularly

testing methods in organizations just like Münsterberg had suggested (Vinchur & Koppes, 2010; Viteles, 1923). Psychotechnology as a movement also catched over to the Netherlands, where the the lab at University of Groningen conducted early testing for selection of telephonists and high-currency workers and the city of Amsterdam opened one of the first psychotechnics laboratories (Haas, 1995; van Strien & Dane, 2001), and Switzerland, where Edouard Claparède founded the *Association Internationale de Psychotechnique* (later the *International Association of Applied Psychology*) in 1920 (Roland-Lévy, 2020).

The most well-known of the early psychotechnology/applied psychology institutes was Stern and Lipmann's *Institut für Angewandte Psychologie und psychologische Sammelforschung* in Berlin. Stern and Lipman implemented one of the first systematic approaches for vocational consulting at the Weimar-republic department of labor in the late 1920s and early 1930s. Stern and Lipmann also founded the "Zeitschrift für Angewandte Psychologie und psychologische Sammelforschung" in 1908 (known as the „Zeitschrift für Angewandte Psychologie from 1916 onward). The journal existed until 1933 when it was banned by the newly appointed Nazi regime. The journal published many early papers on the application and development of tests and preceded the American *Journal of Applied Psychology* by almost a decade even though the contents of the journal did not only include topics related to testing. Stern and Lipmann also initiated an influential book series with the title *Schriften zur Psychologie der Berufseignung und des Wirtschaftslebens* (Papers on the Psychology of Vocational Aptitude and Industrial Efficiency). This paper series had almost the character of a journal and included many early papers on the use of testing in applied settings (see Figure 4) and the young field of vocational testing (e.g., Lipmann, 1921). Contributor to the series were also two of the first widely known female applied psychologists, Hildegard Grünbaum-Sachs and Franziska Baumgarten. Baumgarten later also wrote several books on applied psychology and especially on testing that were widely circulated in the 1930s (Daub, 2011; Richebächer, 2010) including the seminal book *Die Berufseignungsprüfungen: Theorie und Praxis* (Vocational Testing: Theory and Practice; Baumgarten, 1928). Another well-known early applied psychologist was Walter Moede. Moede was a former assistant of Wundt (Moede, 1919, 1920) and studied selection. He was also an early pioneer of the assessment of teams and contributed to the emergence of social psychology.

The Exodus From Germany

Many leading figures in German and Austrian psychology of the late 1920s and early 1930s including Stern and Lipman were Jewish or were critical of the right-wing parties and the general political climate that started to emerge in the Germany of the 1920s. One key element of the climate of the time was the so-called "Dolchstosslegende". According to this widely circulated story the German army was never defeated in the field and was actually close to winning World War I. The Nazis finally took over power in Germany in 1933 with the election of Hitler to the chancellor position (equivalent of a prime minister). This change ended the career of many leading figure in psychological testing with a Jewish background but also created many new opportunities for others like Moede who was a member of the Nazi party.

The victims of the Nazi regime among psychologists include not only psychologists interested in testing and measurement like William Stern, Otto Lipman, but also important figures of the broader discipline of psychology like Wolfgang Köhler, Max Wertheimer, Kurt Lewin, or Otto Selz (Lück, 2011; Sprung & Sprung, 2001). A large number of these important figures left Germany and Austria in the 1930s and mostly found a new home in the United States. However, some scholars like Otto Selz were directly killed by the Nazi regime (Otto Selz died in a concentration camp). The Lipmann-Stern era in Berlin and Hamburg found an infamous end when students who were members of the Nazi organization SA (Sturmabteilung) lay waste

on Lipmann's Berlin institute. Lipmann decided to end his life (Lutz von Rosenstiel, 1985). Shortly thereafter, Stern emigrated to the US via the Netherlands and taught at Duke university until his death. Past his prime, he started his first lecture in the US with the words "First: I realize that my English is imperfect ... This is the first English lecture I have ever given, and so you will have not only instruction, but also occasion for amusement […]." (Probst, 2014).

Wehrmachtpsychologie Testing

The new regime fostered psychological research that seemed closer to the new Nazi ideology and also related disciplines like most notably heredity research that were later heavily associated with war crimes like the work of Otmar Freiherr von Verschuer at the Kaiser Wilhelm Institute of Anthropology focusing on eugenics and human heredity. Within psychology, the change was mostly gradual. Psychologists still endorsed the experimental-scientific ideas of psychotechnology but tried to combine them with theoretical framework known as *Ausdruckspsychologie* (expression psychology) and *Charakterologie* (characterology). Ausdruckspsychologie was developed from the start of the 20th century from graphology and largely was based on the idea that careful and observation of nonverbal behavior would allow inferences about an individuals' personality characteristics (Haas, 1995; Manteufel, 2006). Ausdruckspsychologie-ideas had links to Gestalt psychologists like Karl Bühler (1933) and the notion that a person should be judged "as a whole" (Fitts, 1946; Highhouse, 2002). However, during the time of the Nazi regime, Ausdruckspsychologie was also linked with physiogonomy-related ideas (the practice of assessing a person's character or personality from their outer appearance) and associated with psychologists like Ludwig Klages, who was openly antisemitic, and Philipp Lersch, who enthusiastically supported Hitler and the Nazi euthanasia programs (Ellgring, 1981; Geuter, 1985).

The exodus of leading figures in German vocational testing because of their Jewish heritage or opposition to the Nazi regime opened up space for another group of psychologists. During the years of the Weimar republic, the newly founded army of the new republic faced selection problems. The Treaty of Versailles had limited the number of soldiers in the army considerably so that systematic selection of officers was necessary. The world economic crisis emphasized this trend even more so that even the selection of soldiers required selection procedures. To address this need, the army of the Weimar republic started special military psychology unit that was led by Max Simoneit from the 1930s and exclusively included male psychologists (Geuter, 2008). The unit saw massive growth during the 1930s and conducted careful selection examinations of all officers and later also with specialists of lower ranks. Simoneit was initially not a member of the Nazi party when he was appointed (Bönner, 1986; Lück, 2015). He wrote several books on the psychology of soldiers. Today, he is possibly most well-known for extending and further developing the heterogenous batteries of psychometric tests and practical tasks under observation both individually and in groups used by the Wehrmachtspsychologie. This battery later provided the inspiration for test batteries developed and used by the British War Officer Selection Board and the US Office of Strategic Services (p. 3; The OSS Assessment staff, 1948) and subsequently the basis for the use of similar methods in business organizations after the war where they developed into what is today known as the *assessment center* (Wiggins, 1973).

As the war progressed, the Wehrmachtpsychologie lost a lot of its influence during 1942 because of the need to recruit rather than select in the changed circumstances of the war and the number of tests markedly decreased during 1942 (Geuter, 2008). In this time period, Simoneit joined the Nazi party—possibly in an effort to improve the standing of the organization he lead

(Lück, 2015). Nevertheless, the Wehrmachtspsychologie was ultimately dissolved at the end of 1942 and Simoneit became a soldier and fought on the east front (Geuter, 1985; Lück, 2015). After the war, former members of the Wehrmachtpsychologie took over many of the vacant faculty positions in postwar Germany so the ideas of the former Wehrmachtspsychologie had a latent influence on German psychology into the 1960s and 1970s (Bönner, 1986; de Lorent, 2017). Simoneit himself did not get a position and instead worked as an independent scholar/psychologist after the war (Bönner, 1986; Lück, 2015).

Dutch Psychometrics and CITO

Before and during World War II, Dutch interest in assessment and testing was largely restricted to efforts by individual researchers in the Groningen lab and in industry to apply psychotechnology to practice (Haas, 1995; van Strien & Dane, 2001). The majority of Dutch university psychology had largely been influenced by sociological thinking and testing rarely went beyond a Dutch translation of Binet-styled intelligence testing and its pragmatic application (Mulder & Heyting, 1998). This situation broadly changed in the 1950s when Dutch psychologists became interested in the methods that American testing researchers had developed during the 1930s (Heijden & Sijtsma, 1996). One important figure in the early development of testing in the Netherlands was A. de Groot—a professor at the University of Amsterdam who was a former head of the psychology section of the psychology section at the Philipps company. De Groot founded the Research Instituut voor de Toegepaste Psychologie (RITP) in 1957 where he developed tests for education and selection. De Groot visited the Educational Testing Service in 1958 and tried to implement ideas and technologies developed at ETS in the Netherlands. His school tests were widely used from the 1950s and his efforts culminated in the founding of the 'Centraal Instituut voor Toetsontwikkeling' (CITO; National Institute for Educational Measurement) at Arnhem in 1968 as a central body for the development of tests in the Netherlands.

Dutch testing psychology from early on was interested in item-response and Rasch methods that developed after the initial publication of the Rasch model by the Dane Georg Rasch in 1960. An important figure in this movement was R. J. Mokken—a political scientist by training who worked on the application of IRT methods at the University of Amsterdam. Another important figure in the development of testing in the Netherlands was J. P. van de Geer—a professor at the University of Leiden who organized several meetings on the new methods of Rasch and IRT measurement in the 1960s and 1970s and created a climate that was fruitful for young psychometricians and which led to the success of many well-known Dutch testing and psychometric scholars like De Leeuw, Wim Van der Linden, Ten Berge, Molenaar, Sijtsma and Meijer (Heijden & Sijtsma, 1996).

Internationalization and Critical Psychology

The years after the second world war were characterized by the adoption of US American assessment instruments for clinical, organizational, and educational practice in especially the Dutch-speaking part of Central Western Europe and increasing internationalization (Haas, 1995; van Strien & Dane, 2001). In German-speaking part, a similar adoption occurred to a lesser degree and with some delay (Pawlik, 1994; Sprung & Sprung, 2001). Two developments hampered internationalization. First, many of the former Wehrmachtpsychologen were accepted into influential university positions, continued to focus on Ausdruckspsychologie after the war and into the early 1970s, and affected the policies of the German society of psychology (that also includes members from Austria and Switzerland). A second development that lead to clashes within the German-speaking world was the emergence of the critical psychology movement. The

early West German state still had many former Nazis in important positions and especially young students rebelled against this situation which ultimately fueled the West German student protests of the late 1960s and early 1970s. Critical psychology partly fulfilled a desire for questioning the status quo. The core idea behind the critical psychology movement led by Klaus Holzkamp was a critique of the third-person perspective of "traditional" psychology, its focus on method over content, and the neglect of the role of society and the context in especially the development of clinical "disorders" (Motzkau & Schraube, 2015; Teo, 1998). The goal of proponents of critical psychology was to substitute critical psychology with a psychology from the perspective of the subject. One element of critical psychology was a critique of tests as a method in daily clinical practice because of the artificial nature of test situations and the tendency of tests to be removed from the specific situation of the individual (Rexelius, 1988).

In the French-speaking part of central Europe, internationalization occurred even slower than in the German-speaking part. As noted previously, the tests that Binet had developed were widely used but surprisingly not very frequently in France in his time. However, this trend continued after world war II and until today to some degree. In addition to cultural differences with other countries like the US (Carson, 2007), one reason for the low popularity of Binet's work was likely his early death and the comparably greater attention to more fundamental researchers like Piaget (Lautrey & Ribaupierre, 2004; Schneider, 1992). Some authors have also suggested that Binet had few followers because of how he interacted with his coworkers (Schneider, 1992). Instead of a focus on testing and assessment, French psychology and especially clinical psychology has long been influenced by psychoanalytic work and especially the psychoanalytic theories of the Frenchmen Jacques Lacan (Botbol & Gourbil, 2018; Hook, 2017). Lacan had critical views on behavioral methods and also psychoanalytic work in the ego-psychology tradition that is popular in the US. During the student protests of the late 1960s that also occurred in France, tests were also frequently criticized by left-leaning intellectuals in a somewhat similar way like German critical psychology did (Lautrey & Ribaupierre, 2004).

University-Entry selection in Belgium, Germany, and the Netherlands

The Dutch CITO example did not necessarily broadly translate to the surrounding countries. However, one notable exception is traditionally the selection of medical students. At the start of this chapter, the situation of testing in France was already discussed and the fact that a largely standardized education system typically makes the broad use of tests for the entry to university and the military less of a necessity when exam grades from these school exams can be used for selection (Carson, 2007). France and Germany have traditionally used this system. Validity studies also typically find that selection tests do not predict much additional variance beyond these central exams (Schuler & Hell, 2008). Other central European countries refrain from explicit selection in large parts like Belgium and Austria and instead use the first year of university education to select students. Although costly, the approach typically leads to broader acceptance in society and the common hope is that it increases fairness especially when paired with non-existing or low student fees. One notable exception to the broad absence of tests in university selection, however, has traditionally been medical education because of the high costs of educating medical professionals. The need for selection medical education has led to the development of specific selection tests. The largest of these programs is likely the German Test für medizinische Studiengänge (TMS) that is also used in Austria and Switzerland in somewhat adapted form. The TMS was developed in the 1970s. Because of declining interest in medical education the test fell out of favor for a time but since 2007 it has experienced a resurgence and is now very widely used again in combination with school grades and may be the largest

European selection program with more than 15,000 yearly participants (Chenot, 2009; Trost et al., 1998; Zimmerhofer & Trost, 2008). Belgium also successfully uses a centralized test for the selection of medical students (Lievens, Ones, & Dilchert, 2009).

The specific characteristics of central European education systems notwithstanding the globalization of education, and the fact that a larger percentage of the population seeks to attend university has recently lead to an increased interest in the use of selection tests to select international students and graduate students. Traditionally, most central European university education systems did not distinguish between bachelor and master-level so that a selection from undergraduate to master-level graduate education was not necessary. This situation has recently changed with the Bologna reforms also broadly separating European university education into an undergraduate and graduate part and there is new interest in the use of tests for graduate student selection (Schwager, Hülsheger, Bridgeman, & Lang, 2015).

Modern Use of Tests in Organizations

In more recent years, the use of tests in governmental and business organizations in central Europe shows some variation across countries and industries (Lievens, 2007) and with the more commonly studied US context. However, overall these variations are typically smaller than many applied psychologists may expect. One important difference between central Europe and other parts of the world has long been that the legal situation made the use of inappropriate selection devices and procedures much less of a risk than, for instance, in the US. This situation has traditionally opened up the markets for non-scientific procedures (Kanning, 2010). One reason for this situation is that central European laws typically penalize the loss of immaterial goods in civil law not very highly. Accordingly, the risk in using an inappropriate selection procedure for an organization is typically lower than in countries with UK/US style legal systems. Additionally, countries in central Europe—with the exception of the Netherlands and France to some degree—mostly do not have easily identifiable racial minorities and instead have cultural, linguistic, and religious minorities making it more complex to identify and systematically study minority differences. An exception was a debate in Germany after the reunification when researchers discovered that East-German test participants in personnel selection typically had lower scores on certain measures (Kersting, 1995).

Overall, the legal and political situation in western Europe leaves the decision to implement evidence-based selection programs more strongly to the initiative of specific organizations and decision makers. One example for such an initiative is the use of Thurstonian-inspired selection tests in the selection of German governmental employees. This research program was initiated by a newly founded institute after the wear—the German Society for Personnel Selection (DGP; Deutsche Gesellschaft für Personalwesen) and its head Adolf Otto Jäger and resulted in the development of the Wilde Intelligence Test (Althoff & Jäger, 1981). This measure was used more than 500,000 times in the 1950 and 1960s and is longer and more comprehensive than most other measures of intelligence (Lang, Kersting, Hülsheger, & Lang, 2010). Jäger later became a professor at the University of Berlin and focused on more fundamental intelligence research. However, a revised version of his more modern test—the BIS-r-DGP is still widely used in selection in Germany (Beauducel & Kersting, 2002).

Another initiative with the goal to improve the use of tests in Western Europe is the development of testing quality criteria systems in the Netherlands and Germany (Evers, Sijtsma, Lucassen, & Meijer, 2010; Hagemeister, Kersting, & Stemmler, 2012; Hornke & Kersting, 2004; Kersting, 2008). These evaluation systems go somewhat beyond the US standards for testing by not only providing a set of guidelines but also an institutional structure that supports the

implementation of these guidelines. The German DIN33430 system builds on the infrastructure of the widely known institute for normation that focuses on norms for technical products or services across a large range. The DIN33430 includes the requirement for special training and focuses on the certification of people. People who want to get certified simply need to pass an exam. Psychology departments increasingly make an effort to teach the key skills for passing this exam already in their curriculum. An initiative related to the DIN33430 is the TBS-TK system (Diagnostik- und Testkuratorium der Föderation Deutscher Psychologenvereinigungen, 2018). TBS-TK focuses on tests (and not people). Test are reviewed by two independent reviewers who write a report together. The test reviews are then published by a leading German psychology journal. Like TBS-TK, the Dutch COTAN system focuses on evaluation tests (Evers et al., 2010). While TBS-TK is based on evaluations by reviewers, COTAN focuses more on threshold-based guidelines. Like the DIN33430 system, COTAN also publishes test evaluations of Dutch-speaking tests.

Outlook and Trends

Our review of key episodes in the development of testing in Central Western Europe has shown that many key innovations came from this region of the world until the middle of the last century. We conclude this chapter with some speculation about future developments in selection and testing specific to this part of the world.

One recent trend is the increasing use of testing and assessment for quality control in education especially in the Netherlands and Germany so that psychometric methods have a more indirect role in improving education and higher education outcomes. Universities in Central Europe are increasingly interested in recruiting international students and this creates a new need to assess foreign applicants for Western European universities and organizations and there is also an increasing trend for students within Europe to study in other countries. A similar trend toward globalization also exists in business organizations and this trend creates and increasing need to use testing to compare applicants and employees across countries and language barriers. We expect that these trends will continue.

A second trend affects the education and training of psychologists as one major source of psychometric knowledge and expertise in testing. Traditionally, the education of psychologists in Central Western Europe was relatively institutionalized with the majority of training organized by state universities on the basis of a relatively standardized curriculum and a 4-5 year education to become a psychologist. The Bologna reforms with the introduction of a master/bachelor system have changed this general makeup. There is a trend for shorter and more specialized educations with a focus on specific subdisciplines of psychology like clinical or work and organizational psychology. At the same time, the number of students in most countries have increased. An important question for the future is how this changed landscape with a larger number of psychologists but with a shorter and more specialized education will change the future of psychological testing in applied settings. The European Federation of Psychologists Associations has tried to address this trend by establishing a European quality standard for psychologists, the EuroPsy (European Federation of Psychologists Associations, 2013). However, this initiative has so far not yet had a profound impact because the labor markets and quality standard are still largely focused on the national level.

References

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York, NY: Henry Holt.

Althoff, K., & Jäger, A. O. (1981). Zur Struktur des Wilde-Intelligenz-Tests (WIT) und zu seiner Stellung im Strukturmodell von Jäger 1967. *Diagnostica*, *27*, 215–226.

Asendorpf, J. (1999). *Psychologie der Persönlichkeit* (2nd ed.). Heidelberg, Germany: Springer.
Baumgarten, F. (1928). *Die Berufseignungsrüfungen: Theorie und Praxis*. Bern, Switzerland: A. Francke.
Beauducel, A., & Kersting, M. (2002). Fluid and Crystallized Intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment*, *18*(2), 97–112. https://doi.org/10.1027//1015-5759.18.2.97
Benjamin, L. T. jr. (2006). Hugo Münsterberg's attack on the application of scientific psychology. *Journal of Applied Psychology*, *91*(2), 414–425. https://doi.org/10.1037/0021-9010.91.2.414
Benjamin, L. T. jr. (2009). The birth of American intelligence testing. *Monitor on Psychology*, *40*(1), 20.
Binet, A. (1894). *Introduction à la psychologie expérimentale*. Paris, France: Alcan.
Binet, A., & Simon, T. (1907). *Les enfants anormaux*. Paris, France: Armand Colin.
Blatter, J. T. (2014). The Psychotechnics of Everyday Life: Hugo Münsterberg and the Politics of Applied Psychology, 1887-1917. *ProQuest Dissertations and Theses*, 191. Retrieved from https://search.proquest.com/docview/1557761517
Bönner, K. H. (1986). *Das Leben des Dr. Max Simoneit*. https://doi.org/10.23668/psycharchives.265
Botbol, M., & Gourbil, A. (2018). The place of psychoanalysis in French psychiatry. *BJPsych International*, *15*(1), 3–5. https://doi.org/10.1192/bji.2017.3
Bühler, K. (1933). *Ausdruckstheorie: Das System an der Geschichte aufgezeigt*. Jena, Germany: Gustav Fischer.
Busato, V., van Essen, M., & Koops, W. (2013). *Vier grondleggers van de psychologie*. Amsterdam, The Netherlands: Bert Bakker.
Carson, J. (2007). *The measure of merit: Talents, intelligence, and inequality in the French and American republics, 1750–1940*. Princeton, NJ, US: Princeton University Press.
Cattell, J. M. (1886a). *Psychometrische Untersuchungen*. Leipzig, Germany: Wilhelm Engelmann.
Cattell, J. M. (1886b). Psychometrische Untersuchungen. *Philosophische Studien: Erste Abtheilung*, *3*, 305–335.
Chenot, J.-F. (2009). Undergraduate medical education in Germany. *German Medical Science : GMS e-Journal*, *7*, Doc02. https://doi.org/10.3205/000061
Daub, E. (2011). Franziska Baumgarten-Tramer: Für die Wissenschaftlichkeit praktischer Psychologie. In S. Volkmann-Raue & H. E. Lück (Eds.), *Bedeutende Psychologinnen des 20. Jahrhunderts* (2nd ed.). Wiesbaden, Germany: VS Verlag.
de Lorent, H.-P. (2017). Peter R. Hofstätter. In *Täterprofile: Die Verantwortlichen im Hamburger Bildungswesen unterm Hakenkreuz und in der Zeit nach 1945, Band 2* (pp. 742–797). Hamburg, Germany: Landeszentrale für politische Bildung Hamburg.
Diagnostik- und Testkuratorium der Föderation Deutscher Psychologenvereinigungen. (2018). Föderation Deutscher Psychologenvereinigungen. *Psychologische Rundschau*, *69*(2), 109–116. https://doi.org/10.1026/0033-3042/a000401
Ellgring, H. (1981). Nonverbal communication: A review of research in Germany. *German Journal of Psychology*, *5*, 59–84.
European Federation of Psychologists Associations. (2013). *EuroPsy: European certificate in psychology*. Brussels, Belgium: Author.
Evers, A., Sijtsma, K., Lucassen, W., & Meijer, R. R. (2010). The Dutch Review Process for

Evaluating the Quality of Psychological Tests: History, Procedure, and Results. *International Journal of Testing*, *10*(4), 295–317. https://doi.org/10.1080/15305058.2010.518325
Fahrenberg, J. (2012). *Wilhelm Wundt - Pioneer in psychology and outsider?* Retrieved from http://psydok.psycharchives.de/jspui/handle/20.500.11780/669
Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York: W.W. Norton & Company.
Fitts, P. M. (1946). German applied psychology during World War II. *American Psychologist*, *1*(5), 151–161. https://doi.org/https://doi.org/10.1037/h0059674
Geuter, U. (1985). Polemos panton pater - Militär und Psychologie im Deutschen Reich 1914-1945. In M. G. Ash & U. Geuter (Eds.), *Geschichte der deutschen Psychologie im 20. Jahrhundert* (pp. 146–171). Opladen, Germany: Westdeutscher Verlag.
Geuter, U. (2008). *The professionalization of psychology in Nazi Germany*. Cambridge, UK: Cambridge University Press.
Gundlach, H. (1996). Faktor Mensch im Krieg Der Eintritt der Psychologie und Psychotechnik in den Krieg. *Berichte Zur Wissenschaftsgeschichte*, *19*(2–3), 131–143. https://doi.org/10.1002/bewi.19960190211
Gundlach, H. (2002). *Psychologie und Psychotechnik bei der Eisenbahn*. Retrieved from https://psydok.psycharchives.de/jspui/bitstream/20.500.11780/3466/1/gundlach_01.pdf
Gundlach, H. (2018). William James and the Heidelberg fiasco. *History of Psychology*, *21*(1), 47–72. https://doi.org/10.1037/hop0000083
Haas, E. (1995). *De opkomst van de bedrijfs- en schoolpsychologische beroepspraktijk in Nederland*. Hilversum, The Netherlands: Uitgeverij Verloren.
Hagemeister, C., Kersting, M., & Stemmler, G. (2012). Test Reviewing in Germany. *International Journal of Testing*, *12*(2), 185–194. https://doi.org/10.1080/15305058.2012.657922
Heijden, P. G. M., & Sijtsma, K. (1996). Fifty years of measurement and scaling in the Dutch social sciences. *Statistica Neerlandica*, *50*(1), 111–135. https://doi.org/10.1111/j.1467-9574.1996.tb01483.x
Highhouse, S. (2002). Assessing the candidate as a whole: A historical and critical analysis of individual psychological assessment for personnel decision making. *Personnel Psychology*, *55*(2), 363–396. https://doi.org/10.1111/j.1744-6570.2002.tb00114.x
Hook, D. (2017). The subject of psychology: A Lacanian critique. *Social and Personality Psychology Compass*, *11*(5), e12316. https://doi.org/10.1111/spc3.12316
Hornke, L. F., & Kersting, M. (2004). "Checkliste" zur DIN33430. In L. F. Hornke & U. Winterfeld (Eds.), *Eigenungsbeurteilungen auf dem Prüfstand: DIN33430 zur Qualitätssicherung* (pp. 273–324). Retrieved from http://eprints.uanl.mx/5481/1/1020149995.PDF
Kanning, U. P. (2010). *Von Schädeldeutern und anderen Scharlatanen: Unseriöse Methoden der Psychodiagnostik*. Lengerich, Germany: Pabst.
Kersting, M. (1995). Det Einsatz "westdeutscher" Tests zur Personalauswahl in den neuen Bundesländern und die Fairnessfrage: Auswirkungen der Testleistungsdisparität zwischen Ost und West auf die Auswahlentscheidung. *Report Psychologie*, *20*, 32–41.
Kersting, M. (2008). *Qualität in der Diagnostik und Personalauswahl - der DIN-Ansatz*. Retrieved from http://www.gbv.de/dms/zbw/568932684.pdf%5Cnhttp://d-nb.info/988964112/04

Lang, J. W. B., Kersting, M., Hülsheger, U. R., & Lang, J. (2010). General mental ability, narrower cognitive abilities, and job performance: The perspective of the nested-factors model of cognitive abilities. *Personnel Psychology*, 595–640. https://doi.org/10.1111/j.1744-6570.2010.01182.x

Lautrey, J., & Ribaupierre, A. de. (2004). Psychology of Human Intelligence in France and French-Speaking Switzerland. In R. J. Sternberg (Ed.), *International Handbook of Intelligence* (pp. 104–134). https://doi.org/10.1017/CBO9780511616648.005

Lievens, F. (2007). Research on selection in an international context: Current status and future directions. In M. M. Harris (Ed.), *Handbook of Research in International Human Resource Management* (pp. 107–123). Mahwah, NJ: Erlbaum.

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, *94*, 1514–1535. https://doi.org/10.1037/a0016137

Lipmann, O. (1921). Wirtschaftspsychologie und psychologische Berufsberatung. *Schriften Zur Psychologie Der Berufseignung Und Des Wirtschaftslebens*, *1*(1), 1–24.

Lück, H. E. (2011). Anfänge der Wirtschaftspsychologie bei Kurt Lewin. *Gestalt Theory*, *33*(2), 91–114.

Lück, H. E. (2015). Simoneit, Max. In U. Wolfradt, E. Billmann-Mahecha, & A. Stock (Eds.), *Deutschsprachige Psychologinnen und Psychologen 1933-1945. Ein Personenlexikon, ergänzt um einen Text von Erich Stern* (pp. 418–419). Wiesbaden, Germany: Springer.

Lutz von Rosenstiel. (1985). Lipmann, Otto. In *Neue Deutsche Biographie 14* (p. 645). Retrieved from https://www.deutsche-biographie.de/downloadPDF?url=sfz51743.pdf

Manteufel, A. (2006). "Von den Möglichkeiten einer psychologischen Menschenauslese" - Psychodiagnostik im Kriegsdienst der deutschen Wehrmacht. In G. Lamberti (Ed.), *Intelligenz auf dem Prüfstand: 100 Jahre Psychometrie*. Göttingen, Germany: Vandenhoeck & Ruprecht.

Moede, W. (1919). *Die Experimentalpsychologie im Dienste des Wirtschaftslebens*. Berlin, Germany: Julius Springer.

Moede, W. (1920). *Experimentelle Massenpsychologie*. Leipzig, Germany: S. Hirzel.

Motzkau, J., & Schraube, E. (2015). Kritische Psychologie: Psychology from the standpoint of thesubject. In *Handbook of Critical Psychology* (pp. 280–289). New York, NY, US: Routledge.

Mulder, E., & Heyting, F. (1998). The Dutch curve: The introduction and reception of intelligence testing in the Netherlands, 1908–1940. *Journal of the History of the Behavioral Sciences*, *34*(4), 349–366. https://doi.org/10.1002/(sici)1520-6696(199823)34:4<349::aid-jhbs1>3.0.co;2-m

Münsterberg, H. (1912). *Psychologie und Wirtschaftsleben: Ein Beitrag zur angewandten Experimental-Psychologie*. Leipzig, Germany: Barth.

Münsterberg, H. (1913). *Psychology and industrial efficency*. Cambridge, MA: The Riverside Press.

Nicolas, S. (2006). The importance of instrument makers for the development of experimental psychology: The case of Alfred Binet at the Sorbonne Laboratory. *Journal of the History of the Behavioral Sciences*, *42*(JULY), 3–18. https://doi.org/10.1002/jhbs.21790

Nicolas, S., Andrieu, B., Sanitioso, R. B., Vincent, R., & Murray, D. J. (2015). Alfred Binet and Crépieux-Jamin: Can intelligence be measured scientifically by graphology? *Annee Psychologique*, *115*(1), 3–52. https://doi.org/10.4074/S0003503315001013

Nicolas, S., & Young, J. L. (2014). A French description of the psychology laboratory of G. S. Hall at Clark university in 1893. *American Journal of Psychology*, *127*(4), 527–535. https://doi.org/10.5406/amerjpsyc.127.4.0527

Pawlik, K. (1994). Psychology in Europe: Origins and Development of Psychology in German-Speaking Countries. *International Journal of Psychology*, *29*(6), 677–694. https://doi.org/10.1080/00207599408246558

Probst, P. (2014). „Um den Bedürfnissen des praktischen Lebens entgegenzukommen“ – ein Einblick in Biografie und Werk William Sterns. In M. Spieß (Ed.), *100 Jahre akademische Psychologie in Hamburg. Eine Festschrift* (pp. 87–115). Hamburg, Germany: Hamburg University Press.

Rexelius, G. (Ed.). (1988). *Psychologie als Gesellschaftswissenschaft: Geschichte, Theorie und Praxis kritischer Psychologie*. Opladen, West Germany: Westdeutscher Verlag.

Richebächer, S. (2010). Franziska Baumgarten (1883–1970): Psychotechnikerin, Pionierin, Professorin, Pazifistin. *Punktum*, 5–7.

Richelle, M., Janssen, P., & Bredart, S. (1992). Psychology in Belgium. *Annual Review of Psychology*, *43*(1), 505–529. https://doi.org/10.1146/annurev.ps.43.020192.002445

Roland-Lévy, C. (2020). The first hundred years of IAAP towards the future. *IAAP Bulletin*, *2*(1), 9–10. Retrieved from https://iaapsy.org/site/assets/files/1082/apaw_2020_jan_vol2_1_-_final.pdf

Schneider, W. H. (1992). After Binet: French intelligence testing, 1900–1950. *Journal of the History of the Behavioral Sciences*, *28*(2), 111–132. https://doi.org/10.1002/1520-6696(199204)28:2<111::AID-JHBS2300280202>3.0.CO;2-W

Schuler, H., & Hell, B. (2008). *Studierendenauswahl und Studienentscheidung*. Göttingen, Germany: Hogrefe.

Schumann, F. (2004). *Bericht über den 1. Kongress für experimentelle Psychologie in Gießen vom 18. bis 21. April 1904: Wiederabdruck der Erstausgabe Leipzig 1904 im Auftrag der Deutschen Gesellschaft für Psychologie*. Göttingen, Germany: Hogrefe.

Schwager, I. T. L., Hülsheger, U. R., Bridgeman, B., & Lang, J. W. B. (2015). Graduate Student Selection: Graduate record examination, socioeconomic status, and undergraduate grade point average as predictors of study success in a western European University. *International Journal of Selection and Assessment*, *23*(1), 71–79. https://doi.org/10.1111/ijsa.12096

Siegler, R. S. (1992). The other Alfred Binet. *Developmental Psychology*, *28*(2), 179–190. https://doi.org/10.1037/0012-1649.28.2.179

Spearman. (1904). “ General intelligence ,” objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–292.

Spillmann, J., & Spillmann, L. (1993). The rise and fall of Hugo Münsterberg. *Journal of the History of the Behavioral Sciences*, *29*(4), 322–338. https://doi.org/10.1002/1520-6696(199310)29:4<322::AID-JHBS2300290403>3.0.CO;2-1

Sprung, L., & Sprung, H. (2001). History of modern psychology in Germany in 19th- and 20th-century thought and society. *International Journal of Psychology*, *36*(6), 364–376. https://doi.org/10.1080/00207590143000199

Stern, W. (1900). *Über Psychologie der individuellen Differenzen: Ideen zu einer 'differentiellen Psychologie'*. Leipzig, Germany: Barth.

Stern, W. (1911). *Die differentielle Psychologie in ihren methodischen Grundlagen*. Leipzig, Germany: Barth.

Stern, W. (1914). *Die psychologischen Methoden der Intelligenzprüfung: und deren Anwendung*

an Schulkindern. Leipzig, Germany: Barth.
Stern, W. (1917). Hugo Munsterberg: In memoriam. *Journal of Applied Psychology*, *1*(2), 186–188. https://doi.org/10.1037/h0074508
Stout, S. E. (2018). Training Soldiers for the Roman Legion. *The Classical Journal*, *16*(7), 423–431.
Teo, T. (1998). Klaus Holzkamp and the rise and decline of German critical psychology. *History of Psychology*, *1*(3), 235–253. https://doi.org/10.1037/1093-4510.1.3.235
Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Cambridge, MA: Houghton Mifflin.
The OSS Assessment staff. (1948). *Assessment of men: Selection of personnel for the Office of Strategic Services*. New York, NY, US: Rinehart & Company.
Trost, G., Blum, F., Fay, E., Klieme, E., Maichle, U., Meyer, M., & Nauels, H.-U. (1998). *Evaluation des Tests für medizinische Studiengänge (TMS): Synopse der Ergebnisse*. Bonn, Germany: Institut für Test- und Begabungsforschung.
van Strien, P. J., & Dane, J. (2001). *Driekwart eeuw psychotechniek in Nederland: de magie van het testen*. Arnhem, The Netherlands: Nederlandse Stichting voor Psychotechniek.
Vinchur, A. J., & Koppes, L. L. (2007). Early contributions to the science and practice of industrial psychology. In L. L. Koppes, P. W. Thayer, A. J. Vinchur, & E. Salas (Eds.), *Historical perspectives in industrial and organizational psychology* (pp. 37–58). New York, NY: Psychology Press.
Vinchur, A. J., & Koppes, L. L. (2010). A historical survey of research and practice in industrial and organizational psychology. *APA Handbook of Industrial and Organizational Psychology, Vol 1: Building and Developing the Organization.*, 3–36. https://doi.org/10.1037/12169-001
Viteles, M. S. (1923). Psychology in business-in England, France, and Germany. *The Annals of the American Academy of Political and Social Science*, *110*, 207–220. Retrieved from https://www.jstor.org/stable/1015087
Wickett, J. C. (1998). Hans J . Eysenck ' s Influence on Intelligence Research. *Psihologija*, 249–256.
Wiggins, J. S. (1973). *Personality and prediction: Principles of Personality Assessment*. Reading, MA, US: Addison-Wesley.
Wissler, C. (1901). The correlation of mental and physical tests. *The Psychological Review: Monograph Supplements*, *3*(6), i–62. https://doi.org/10.1037/h0092995
Wundt, W. (1883). Über die Messung psychischer Vorgänge. *Philosophische Studien*, 251–260. Retrieved from http://vlp.mpiwg-berlin.mpg.de/
Zimmerhofer, A., & Trost, G. (2008). Auswahl- und Feststellungsverfahren in Deutschland - Vergangenheit, Gegenwart und Zukunft. In H. Schuler & B. Hell (Eds.), *Studierendenaushal und Studienentscheidung* (pp. 32–42). Göttingen, Germany: Hogrefe.