Springer Proceedings in Mathematics & Statistics

Marie Wiberg · Dylan Molenaar
Jorge González · Ulf Böckenholt
Jee-Seon Kim *Editors*

Quantitative Psychology

84th Annual Meeting of the Psychometric Society, Santiago, Chile, 2019

MOREMEDIA

Springer

Springer Proceedings in Mathematics & Statistics

Volume 322

Springer Proceedings in Mathematics & Statistics

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including operation research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at http://www.springer.com/series/10533

Marie Wiberg • Dylan Molenaar • Jorge González
Ulf Böckenholt • Jee-Seon Kim
Editors

Quantitative Psychology

84th Annual Meeting of the Psychometric Society, Santiago, Chile, 2019

Springer

Editors
Marie Wiberg
Department of Statistics, USBE
Umeå University
Umeå, Sweden

Dylan Molenaar
Department of Psychology
University of Amsterdam
Noord-Holland
Amsterdam, The Netherlands

Jorge González
Facultad de Matematicas
Pontificia Universidad Católica de Chile
Santiago, Chile

Ulf Böckenholt
Kellogg School of Management
Northwestern University
Evanston, IL, USA

Jee-Seon Kim
Department of Educational Psychology
University of Wisconsin–Madison
Madison, WI, USA

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-030-43468-7 ISBN 978-3-030-43469-4 (eBook)
https://doi.org/10.1007/978-3-030-43469-4

Mathematics Subject Classification: 62P15, 62P25, 91E45

Preface

This volume represents presentations given at the 84th Annual Meeting of the Psychometric Society, organized by Centro de Extensión at the Pontificia Universidad Católica de Chile, in Santiago, Chile, on July 15–19, 2019. The meeting attracted 411 participants, and 383 papers were presented, of which 84 were part of a symposium. There were 4 preconference workshops, 11 keynote presentations, 8 invited presentations, 2 career-ward presentations, 4 state-of-the-art presentations, 66 poster presentations, 1 dissertation award winner, and 19 symposia.

Since the 77th meeting in Lincoln, Nebraska, Springer has published the proceedings volume from the annual meeting of the Psychometric Society to allow presenters to spread their ideas quickly to the wider research community while still undergoing a thorough review process. The previous seven volumes of the meetings in Lincoln, Arnhem, Madison, Beijing, Asheville, Zurich, and New York were enthusiastically received, and we expect these proceedings to be successful as well.

The authors of these proceedings were asked to use their presentations at the meeting as the bases of their chapters, possibly extended with new ideas or additional information. The result is a selection of 28 state-of-the-art chapters addressing a diverse set of psychometric topics, including but not limited to item response theory, factor analysis, hierarchical models, and computerized adaptive testing.

Umeå, Sweden	Marie Wiberg
Amsterdam, Noord-Holland, The Netherlands	Dylan Molenaar
Santiago, Chile	Jorge González
Evanston, IL, USA	Ulf Böckenholt
Madison, WI, USA	Jee-Seon Kim

Contents

Stories of Successful Careers in Psychometrics and What We Can Learn from Them

Carolyn J. Anderson, Susan Embretson, Jacqueline Meulman, Irini Moustaki, Alina A. von Davier, Marie Wiberg, and Duanli Yan

Abstract This paper was inspired by the presentations and discussions from the panel "Successful Careers in Academia and Industry and What We Can Learn from Them" that took place at the IMPS meeting in 2019. In this paper, we discuss what makes a career successful in academia and industry and we provide examples from the past to the present. We include education and career paths as well as highlights of achievements as researchers and teachers. The paper provides a brief historical context for the representation of women in psychometrics and an insight into strategies for success for publishing, for grant applications and promotion. The authors outline the importance of interdisciplinary work, the inclusive citation

C. J. Anderson (✉)
University of Illinois, Champaign, IL, USA
e-mail: cja@illinois.edu

S. Embretson
Georgia Institute of Technology, Atlanta, GA, USA
e-mail: susan.embretson@psych.gatech.edu

J. Meulman
Leiden University, Leiden, The Netherlands

Stanford University, Stanford, CA, USA
e-mail: jmeulman@math.leidenuniv.nl; jmeulman@stanford.edu

I. Moustaki
London School of Economics & Political Science, London, UK
e-mail: i.moustaki@lse.ac.uk

A. A. von Davier
ACTNext, Iowa City, IA, USA
e-mail: Alina.vonDavier@act.org

M. Wiberg
Department of Statistics, USBE, Umeå University, Umeå, Sweden
e-mail: marie.wiberg@umu.se

D. Yan
Educational Testing Service, Princeton, NJ, USA
e-mail: dyan@ets.org

M. Wiberg et al. (eds.), *Quantitative Psychology*, Springer Proceedings in Mathematics & Statistics 322, https://doi.org/10.1007/978-3-030-43469-4_1

approaches, and visibility of research in academia and industry. The personal stories provide a platform for considering the needs for a supportive work environment for women and for work-life balance. The outcome of these discussions and reflections of the panel members are included in the paper.

Keywords Advice · Career paths · Psychometrics history · Gender gap

1 Introduction

In recent years, society has started to shift its narrative about scientists from the lonely genius (usually a white man) to more diverse images of the researchers, authors of papers, and to their supportive environment. The IMPS19 session, "Stories of Successful Careers in Psychometrics and What We Can Learn from Them," is part of this expansion of acknowledgment of the contributions of contemporary fellow scientists to the field of psychometrics and their individual paths to successful careers. This proceedings volume provides a snapshot of the interests of members of the Psychometric Society in 2019 and as such it encompasses a historical and social perspective on ideas, creators, and life stories that are being mingled with the psychometric papers that these authors published in this volume or elsewhere.

In this paper, we loosely follow the structure of the symposium and allow the contributors to speak to her professional successes and to the personal context in which these successes took shape. The professional successes include breakthrough research ideas and projects, leadership acknowledgment, and social impact. The scientists will also share their lessons learned for the next generations of psychometricians. The team of established scientists is comprised of seven women from six countries, who now live and work across four countries. Some of these stories speak to the geopolitical influence, the immigrant's experience, the struggle to publish in a foreign language, and the struggle to be authentic in a professional world with relatively narrow expectations.

There are many socio-historical, political, and cultural conditions that have led to marginalization of women in technical domains. STEM subjects in some societies are highly gendered often based on a belief that boys are better at math than girls due to biological differences. In the USA, women earn fewer PhDs in STEM domains and only 31.5% of women earned PhDs in mathematics and computer (Okahana and Zhou 2017). School and parental guidance have also contributed to the gender gap in STEM. Girls and boys often grow up with the idea that they will be bad and good at math, respectively (e.g., Math class is tough! Barbie is for girls) and that girls do not belong in a technical environment. All those reasons are in addition to systemic and structural biases such as opportunities for training, and later on, for recruiting.

An article published in *The Guardian* by Carol Black and Asiya Islam in 2014 is a response to over 50 senior Cambridge academics called on the university to change its staff appointment procedure because the existing system favored men. They stated that "Despite accounting for 45% of the academic workforce, women

hold only 20% of professorships in UK universities, and just 15.3% of such posts in Cambridge" (Black and Islam 2014). Though more women enter university than men and there is an almost equal representation of women and men at lower professional levels, only 27.5% of senior managers in higher education and 20.5% of professors in the UK are women. Worse, only 1.1% of senior managers in higher education and 1.4% of professors in the UK are black and minority ethnic women.

One would expect that in more gender-equal societies the gender gap in STEM scores and in higher managerial or academic positions is smaller. Different systems for tenure and promotion also lead to different outcomes. It looks like the problem is universal. This imbalance is spread in the world and the causes are often blatantly attributed to narrow views of women's roles in society: *The Guardian* reported in June 2019, that after a medical school in Japan admitted rigging admission procedures to give men an unfair advantage, once the system became fair, women have outperformed their male counterparts in entrance examinations (McCurry 2019).

2 History

This section provides a brief historical tribute to women who have contributed to psychometrics and related disciplines. It is by no means complete or exhaustive. Psychometrics was founded by Thurstone's vision for a mathematical underpinning for psychological research. The Psychometric Society was founded in 1935 by Louis Thurstone, Jack Dunlap, Paul Horst, Albert Kurtz, Marion Richardson, and John Stalnaker. Paul Horst and Albert Kurtz founded the journal in 1936 with the mission to create a journal that will be mathematically oriented to develop and disseminate work in psychological measurement. Much before that, Gauss in 1809 presented the theory of errors of observation following the normal distribution, Bessel's presented "a personal equation" to correct observations for differences among observers, Galton in 1884 designed an apparatus to measure a variety of bodily dimensions, Cattell in 1889 established a laboratory of psychology with an interest in psychometric measures.

Where are the women in all those initiatives and contributions? At a time when women were destined to get married and bear children, Florence Nightingale (1820–1910), who was self-educated in statistics, pioneered in visual statistical graphs called Nightingale Rose Diagram or Polar Area Diagram. She was the first female member of the Royal Statistical Society and the founder of the nursing profession. Florence Nightingale David (1909–1993), named after Florence Nightingale, studied mathematics at Bedford College for Women after failing to go to University College of London. She published the *Tables of the Correlation Coefficient*, as well as *Combinatorial Chance* (with D.E. Barton) and *Games, Gods and Gambling: The Origins and the History of Probability*. She chaired the statistics department at the University of California, Berkeley, then founded the statistics department at the University of California, Riverside.

Ethel Elderton (1878–1954) is a true hidden figure, a female researcher who worked with Galton and Pearson in eugenics research. In 1905 she resigned her teaching post to become Galton's assistant. Subsequently, she became a Galton Scholar and Fellow and Assistant Professor at University College London. In the same period, Gertrude Mary Cox (1900–1978) dreamed to be a missionary and saving souls in far-off lands. To be qualified as a missionary, she became a student of George Snedecor then published *Experimental Design* (Cochran and Cox 1957). She was the first female department chair in a men's world and started the well-known North Carolina "Research Triangle." In psychometrics, Thelma Thurstone (1897–1993) a psychometrician herself combined the theory of intelligence with its measurement to design instructional materials, like the tests she developed for the American Council on Education from 1924 to 1948. In 1955, Thelma Thurstone was asked to assume the directorship of the Psychometric Laboratory upon the death of her husband in order to continue his funded research projects. Barbara Stoddard Burks (1902–1943) worked in behavioral genetics and intelligence and was the first one who used a graph to represent a mediator. Her first paper published in 1926 was on the inadequacy of the partial and multiple correlation technique. Anne Anastasi (1908–2001) is known as the "test guru" psychometrician and the psychology's female voice. She pioneered the development of psychometrics and chaired the department of psychology at the male-dominated school at Fordham University, and she won many awards including The American Psychological Foundation's Gold Medal for Life Achievement. Her books on *Differential Psychology*, *Fields of Applied Psychology*, and Psychological Testing (with 7 editions) influenced generations of psychometricians. Fordham University established a special position named Anne Anastasi Chair Professor.

Another important contributor is Dorothy Adkins (1912–1975), an American psychologist who was interested in new (at the time) statistical techniques of factor analysis. She applied factor analytic techniques in order to examine and better understand curriculum, program evaluation, and affect in children. She was also co-editor of *Psychometrika* with Paul Horst (1958–1959, 1963–1966) and president of the Psychometric Society in 1949–1950. Forty-five years later Fumiko Samejima became the next female president of the society (1996–1997) who is known for her work on Item Response Theory (IRT) models for polytomous data. A few more women followed as presidents of the society, Susan Embretson (1998–1999), Jacqueline Meulman (2002–2003), Sophia Rabe-Hesketh (2014–2015), and Irini Moustaki (president-elect, 2020–2021). Susan Embretson was the first to integrate cognitive theory into IRT and test design whereas Jacqueline Meulman made significant contributions in the area of multivariate data analysis with optimal transformations of variables, and multidimensional scaling.

Many women without a PhD also made significant contributions. At Harold Gulliksen's Gold Medal Award for Lifetime Achievement in Psychological Science (1991), he acknowledged his wife as his significant collaborator who did the programming and analyses for him. Similarly, Marilyn Wingersky worked mostly with Fred Lord and implemented algorithm, statistical models, and developed the LOGIST software for estimating latent traits and item parameters. Martha Stocking,

without a doctorate, also worked with Fred then furthered her contributions on computerized adaptive testing (CAT) research and development including automated test assembly (ATA) using weighted deviation and the conditional item exposure control algorithm with Charlie Lewis. Kikumi Tatsuoka (1930–2016) received her PhD later in life, after raising her children; she developed the Rule-Space model for diagnostic assessment. Dorothy Thayer has been an instrumental behind the scene figure. She worked with Mel Novick, Don Rubin, Paul Holland, Rebecca Zwick, Charlie Lewis, and Alina von Davier, and published numerous numbers of papers with them, always as the second author. Among other researchers we would like to note is Frances Swineford (1909–1997) who in 1937 together with Holzinger introduced the bifactor model (one general factor and multiple group factors) for mental abilities (Holzinger and Swineford 1937). Again, this oversight has been characteristic of the scientific world in the twentieth century. As discussed in Yong (2019) and Huerta-Sanchez and Rolfs (2019), our colleague professor Margaret Wu from Melbourne has been only thanked for an algorithm that she co-created to compute the "Watterson estimator."

Finally, a very important initiative in 2004 is the Psychology Feminist Voices project directed by Alexandra Rutherford at York University in Toronto, Canada which aims to collect, preserve, and share the narratives of diverse feminist psychologists from all over the world (see http://www.feministvoices.com/about).

3 The Impact of the Structure of Society and Academia

3.1 Societal Structures

The structure of a society is important when pursuing an academic career. To have well-organized paid maternal and paternal leave tend to enhance gender equality. In the past, more men than women earned PhDs, but now in many countries many universities and colleges have more women than men earning PhD degrees (Okahana and Zhou 2017). The balance between work and family has gained attention with both men and women working. Mason and Wolfinger (2013) have examined the relationship between family formation and academic careers of men and women, including an examination of the family sacrifices women often have to make to get ahead in academia and consider how gender and family interact to affect promotion to full professor, salaries, and retirement. Although their research is from the USA it is seen in many countries that even if women and men work a similar number of hours, women tend to take more responsibility for their family. They concluded that men can get a career advantage when having children but for women it can be a career killer. Those women who advance through the faculty ranks tend to pay a high price by being less likely to be married with children. For a woman to facilitate her career it is thus important to be in a relationship which

believes in equality and to work in a country where the society helps women and men with this equality by, for example, paid maternal and paternal leave.

3.2 Academic Structures

It is not just the structure of the society which is important but also the academic structure. To have open calls and transparency in the career system is typically viewed as a way to frame gender equality. van den Brink et al. (2010) examined transparency in the Netherlands and concluded that transparency and accountability should be deployed to their full potential. In their study, transparency was limited to recruitment protocols, but transparency should also imply making the process and decisions more visible for the larger academic society, which is the case in Sweden and Finland.

Internal structures are also important. To be part of a supportive work environment, and to have role models, mentors, and colleagues all greatly enhance the chances of being able to pursue an academic career. Receiving constructive feedback is essential for career development for everyone. However, when Rubini and Menegatti (2014) examined the language in academia, they concluded that judgments of female applicants in academic personnel selection were formulated using negative terms at a more abstract level and positive terms at a more concrete level than those of male applicants. They also found that linguistic discrimination was perpetrated only by male committee members. The discrimination was mainly based on the use of negative adjectives and thus this could be a hindrance for women's academic careers. To counteract this tendency, institutions often try to have men and women represented on different committees; however, women should make sure not to get stuck doing committee work because they need a woman. It is important to say yes to exciting new projects and collaborations and often say no to the role of "female representative" unless you feel they asked you due to your competence. In summary, choose your service and work wisely.

4 Personal Reflections

In this section, each of the panel members has sketched a short biography together with some personal reflections.

4.1 Personal Reflection by Carolyn J. Anderson

Themes throughout CJA's career have included accepting opportunities that were offered to her and following her interests. Curiosity has been a driving force in her

career. In college, CJA was introduced to quantitative psychology by Bill Meredith and Barb Mellers at the University of California at Berkeley and took Bill's graduate seminars on factor analysis and latent class analysis. CJA was hooked!

CJA's first major challenge was choosing a dissertation topic upon which she built a career in academia. The University of Illinois at Urbana-Champaign (UIUC) was an ideal environment to pursue a PhD due to the breadth and depth of expertise of the faculty. Before CJA's ideas solidified, she did research on judgment and decision making with Michael Birnbaum and Elke Weber, and social network analysis with Stanley Wasserman. Stanley agreed to be her advisor and allowed her the freedom and support to pursue and explore whatever interested CJA. Starting with two papers by Leo Goodman that Stanley recommended, CJA read backward, forward, and side-ways in literatures on categorical data analysis, matrix decompositions, graphical models, optimal scaling, and computing algorithms. CJA's dissertation encompassed all of these areas and earned her the Psychometric Society and APA Division 5 Dissertation awards.

Dual career couples can face many challenges, especially finding positions in the same city and having a family. CJA was offered a tenure track position at UIUC and accepted it because she was expecting her first child and both parents would be employed. The policies at UIUC were nonexistent regarding childbirth and family policy. When CJA began, 80% of the tenured faculty in her primary college were men and attitudes of some senior faculty were not supportive of women. For example, after being denied a release from teaching due to childbirth, she was asked "doesn't it bother you that someone else is raising your child?". Fortunately, she also had very supportive colleagues. Stanley Wasserman and Rod McDonald stepped up and taught her courses until she was able to return to work.

After a rocky start, 15 months of little to no sleep, and becoming visually disabled, she needed to jump start her research program. She went back to the literature, including original sources. Typographic errors in a paper had carried through the literature and after correcting them it became obvious that row-column association models and their extensions were standard item response models. This led to an NSF grant and papers on graphical models and latent variables models starting with Anderson and Vermunt (2000).

4.2 Personal Reflection by Susan Embretson

SE's research direction has focused on understanding the cognitive processes, skills, and strategies that are involved in responding to test items. Her research has included developing item response theory models, perspectives on the validity concept, examining the impact of item design on test correlates and developing automatic item generators. SE has received career awards for this research, the 2019 Career Award for Lifetime Achievement from the Psychometric Society, the 2018 Saul Sells Award for Distinguished Multivariate Research from the Society for Multivariate Experimental Psychology, the 2013 Career Contribution Award from

the National Council on Measurement in Education, and the 2011 Distinguished Lifetime Achievement Award from the American Educational Research Association: Assessment and Cognition Division, as well as several scientific contribution awards. Although her personal journey to her research program was not direct, her interests in the topic began in high school after taking the Preliminary Scholastic Aptitude Test. Unfortunately, her score was not high enough to qualify for a Merit Scholarship. Why were test items involving Victorian novel vocabulary on a college admissions test? She complained to her high school counselor and a few weeks later, two individuals from the University of Minnesota came to administer an individual intelligence test. Shortly afterwards she received a full scholarship. Al Johnson, an engineer who built skyscrapers, decided he could fund ten students per year. He probably did not read Victorian novels either.

SE began her studies with a goal to major in psychology. However, the required research experiences in the introductory course, which included running rats in mazes and learning nonsense syllables, did not pique her interest. She changed her major to Spanish, but after learning to speak the language, she found that she was not as enthusiastic about the literature and could not envision being a high school language teacher. By this time, she had a young daughter and a husband, which involved 2 h commuting as they could afford only one car. One very cold winter day, she decided to drop out of school. She hoped that the world of business would suit her better. It did not. After 6 months of the world of work, she decided to take two night school classes: Individual Differences and Psychological Statistics. Wow! SE found her interests. She returned full time to the University of Minnesota and, fortunately, the Al Johnson Foundation decided that they could fund her again and she finished in a little over 1 year.

SE applied for graduate school at the University of Minnesota. Required was a test used to select students for fellowships, the Miller Analogy Test. She remembers the test well. Why is knowing the answer to analogies such as "Moscow: Vodka:: Copenhagen:?" measuring aptitude? She did not know what the Danes drank. Again her score was not high. Despite that, she was selected, primarily because her Bachelor of Arts degree was awarded summa cum laude.

Her graduate career was exciting, as IRT was just entering the field and she was able to pursue her research interests in cognition and measurement. She delayed finishing by 1 year and then took a post doc position for 1 year so that her husband could finish his PhD. Afterwards she interviewed at the University of Georgia. However, the available teaching topic was not her major interest and the work-family balance did not work out. Thus with difficulty, she turned down their offer even though nothing else was pending. This was a good decision, as good luck came in a couple of weeks! The University of Kansas offered her a position to ease into teaching graduate statistics, and she could pursue whatever research topic interested her so long as it was successful. Also, women's expertise in quantitative methods was not questioned, since Julie Shafer had been teaching statistics there. SE accepted the offer, to which she attributed much of her success. She spent 30 years

there and pursued her research interests with enthusiasm. Her current position at the Georgia Institute of Technology has been successful due to the solid base of research and teaching that she built at KU. In summary, SE characterizes her personal journey as involving some good luck, some good decisions, and lots of persistence.

4.3 Personal Reflection by Jacqueline Meulman

After JM was drawn into psychometrics while studying its history, preparing an undergraduate course in History of Psychology as TA at Leiden University, she abandoned everything else by becoming an RA at the Leiden Department of Data Theory in 1978. This department was founded at Leiden University in 1970 by the late John P. van de Geer, and its mission was the development of new and innovative methods for statistical multidimensional data analysis. Later on, Jan de Leeuw added to its mission the implementation in software for multivariate analysis of categorical data, and for multidimensional scaling and unfolding. JM had found the topic in statistics that she would cherish for the next 40+ years to come.

Like Jan de Leeuw and Willem Heiser, JM visited the famous AT&T Bell Telephone Laboratories in Murray Hill, NJ. The year 1982 that she spent in Doug Carroll's group in Mike Wish's department *Computer-Aided Information Systems* changed her life. Doug was a superb mentor who introduced her to all her heroes in psychometrics and beyond. It was Paul Tukey, the nephew of Bell Labs' Associate Executive Director John Tukey, who told her she was not a psychometrician, but a statistician. After returning to the Department of Data Theory in 1983, JM finished her dissertation in 1986 (advisors Jan de Leeuw and John P. van de Geer), and was awarded a 5-year fellowship from the Royal Netherlands Academy of Arts and Sciences, which allowed her to continue her career at the department that she loved.

In 1987, John P. van de Geer retired, and Jan de Leeuw took a position at UCLA, and Willem Heiser and JM were left some big shoes to fill. Their efforts resulted in Albert Gifi's Nonlinear Multivariate Analysis published by Wiley in 1990, and the incorporation of the associated software programs in the SPSS package CATEGORIES (also from 1990 onwards).

A next important period in JM's career started in 1992, by visiting the University of Illinois at Urbana-Champaign, where she was teaching and started collaborating with Larry Hubert. In 1994, JM was awarded the prestigious PIONEER Award by the Netherlands Organization for Scientific Research (NWO), which allowed her to start her own research group in Leiden, as well as spending time in Champaign-Urbana, where she had been appointed as Adjunct Professor in 1993. The collaboration with Larry Hubert and Phipps Arabie (in the so-called HAM team) resulted in a number of papers and two books.

In the meantime, Willem van Zwet, who was Professor of Mathematical Statistics in the Mathematical Institute in Leiden, took it upon him to support JM to become full professor. Her Chair was called Applied Data Theory, and she was leading a group of assistant professors, postdocs, and PhD students; the group was still called

Data Theory, but was relocated at the Department of Education. This association did not develop into a good synergy, and after a number of difficult years, the Data Theory Group left the Department of Education. However, good things also happened in this period: JM was elected as President of the Psychometric Society (in 2001), and as Member of the Royal Netherlands Academy of Arts and Sciences (in 2002).

In 2006, JM was offered a position in statistics at the Leiden Mathematical Institute for one day a week, and this appointment was extended to a full-time position with a Chair in Applied Statistics in 2009. In the meantime, the collaboration with SPSS had resulted in many new software programs, and royalties for Leiden University (first shared with Willem Heiser, and later under full control of JM) that increased to very impressive figures. The latter made it possible for JM to start anew within the Mathematical Institute (MI), with appointing assistant professors and a group of PhD students. At the MI, JM developed with Richard Gill, and later Aad van der Vaart, a new Master program called *Statistical Science (for the Life and Behavioral Sciences)*, in collaboration with other statisticians from the Leiden University Medical Center, the Methodology & Statistics Division at the Leiden Institute of Psychology, and Wageningen University and Research Center. From 2011 to 2016, JM was President of the *Netherlands Society of Statistics and Operations Research*, and she was appointed in the Department of Statistics at Stanford University, first in 2009 as Visiting, and later in 2017 as Adjunct Professor. The above story may sound as a dream, but the path has known many large obstacles, professional as well as medical. JM had to work very hard to pursue her ideals. But all is well that ends well: JM was honored with the Psychometric Society's Career Award for Life Time Achievement 2020.

4.4 Personal Reflection by Irini Moustaki

IM studied Statistics and Computer Science at the Athens University of Economics and Business and continued her studies at the London School of Economics from where she received a masters and PhD in Statistics. Initially, her PhD thesis was on sample surveys and variance estimators under the supervision of Colm O'Muircheartaigh but as soon as Colm was awarded a state grant as a co-investigator with David Bartholomew and Martin Knott on the Analysis of Large and Complex Data Sets, IM started working on latent variable models for mixed data closely also with Knott and Bartholomew. At LSE she has been very fortunate to have had a very supportive and encouraging environment in which to study and later to work. A year before she received her PhD, she got an appointment as a temporary lecturer at LSE and a year later a tenure track position in the same department. The Statistics Department at the time had no female professors and only one female lecturer. IM also spent a period of 5 years at the Athens University of Economics and Business as Assistant and Associate Professor before returning to LSE again in 2007 as associate

professor and in 2013 became a full professor. IM served both as head and deputy head in her department at LSE.

A turning point in her PhD studies was when she attended the IOPS meeting in Tilburg as a PhD student to discover to her surprise a whole community of researchers working on models with latent variables. At LSE and in the UK in general there wasn't much of a psychometric tradition or use of latent variable modeling in social sciences. The second opportunity came when her supervisor encouraged her to attend a workshop by Karl Joreskog on SEM in Heidelberg. This is also the place when she met with Alina and Matthias von Davier and also started a conversation with Karl Joreskog on IRT and SEM that later on led to two papers and a long-term friendship. The Psychometric meetings and community provided her with an academic family which allowed her to discuss her research developments, make collaborators, and make valuable friendships. IM is indebted to the continuous support she received in her early career by Martin Knott at the LSE, who trusted her capabilities and generously exchanged ideas of research and projects. Her collaborations with researchers from LSE but also other places in Europe and beyond led to publications in the areas of missing values, detection of outliers, and composite likelihood estimation. The highlights of her career were when she received an honorary doctorate from the University of Uppsala on the recommendation of her collaborators and friends Fan Wallentin and Karl Joreskog, served as the editor-in-chief of Psychometrika, and honored to be the president-elect of the Psychometric Society. The Psychometric Society has continuously provided a stimulated intellectual environment for her. Further to her teaching and research, IM finds the mentoring of junior academics and PhD students a very important part of her job.

4.5 Personal Reflection by Alina von Davier

AvD studied mathematics at the University of Bucharest and at the end of the studies was fortunate to experience the political change in a country that had been under an authoritarian regime for a long time. The political changes brought opportunities and hope and AvD went to work for a research institute (The Institute of Psychology of the Romanian Academy) instead of teaching math at a high school, as would have been the case under the previous system. Further on, she went to do her PhD in Germany. She started her work on falsifying causal hypotheses with Rolf Steyer, but she discovered interesting singularity points in the testing of hypotheses that captured her interest, and therefore her dissertation ended up back in mathematics, with a second advisor, Norbert Gaffke. In the 5 years she lived in Germany, she also learned German, married MvD, and had a son—efficiently, as she likes to describe it.

The von Daviers moved to the US and specifically to ETS, where their interests found a good match with the company's needs. ETS provided an incredible intellectually rich environment for the development and exploration of one's ideas. Her research journey went from research in test equating, to adaptive testing, and

to the measurement of collaborative problem solving and other complex constructs. She was fortunate to work closely with Paul Holland, Charlie Lewis, and Shelby Haberman. She also became increasingly involved with the operational testing and with the implementation of new methodologies and technologies. In 2015, she introduced the concept of Computational Psychometrics to define the blend of psychometric theory with the data-driven discovery. She moved to ACT in 2016 to establish and lead an innovation hub to help transform the company. With this move, a special opportunity was offered to her to redefine what the educational experience means in the twenty-first century and how psychometrics can be the foundation for the learning, measurement, and navigation efforts to support this experience for everyone everywhere.

4.6 Personal Reflection by Marie Wiberg

MW has in her career been driven by curiosity and she loves to try to solve new challenges and to collaborate with other curious persons. MW started her PhD in Statistics but worked at an educational measurement department where she came into contact with real test problems. From networking at conferences, she ended up as a visiting researcher with Professor Ramsay at McGill University and then moved on to do a postdoc with professor van der Linden at the University of Twente. These two research experiences had a major impact on her future career path. The work with nonparametric item response theory with Ramsay, which they both thought was an "easy" problem to solve, took more than 12 years to solve, but several papers and workshops have followed in recent years. An important lesson is that good ideas and how to solve them may take a while. The work in the Netherlands rerouted her to different test equating problems—a path she still follows and led to successful collaborations with researchers from around the world. Most of her collaborations spring from brief meetings at conferences where many new ideas have emerged. Since the start of her PhD program, MW has had an interest to work with real empirical test data (including national tests, admissions test, and the large-scale assessments TIMSS and PISA). MW recommends everyone who has a chance to work with real data to take the opportunity as many theoretical research problems may emerge. MWs work has been recognized nationally through large research grants and she has been a member of the Young Academy of Sweden which is an academy for talented young researchers within all research fields. Internationally, she has coauthored a test equating book (González and Wiberg 2017), worked as an associate editor for the Journal of Educational Measurement and is currently editor of the IMPS proceedings.

4.7 Personal Reflection by Duanli Yan

DY has been very fortunate to have many distinguished teachers and mentors who have had great influence through the decades on her career and life. DY became interested in statistics and optimization after earning her bachelor's degree in computer science and applications. When she completed her dual masters in statistics and in operations research in the statistics department at Penn State University, Professor C. R. Rao tried to persuade her to stay and do a PhD with him. However, she had been in school for almost all of her life by that time, and she wanted to work. Soon after she started working at ETS, she realized that she should have done a PhD.

While working at ETS, she learned the Rule-space model for cognitive diagnoses from Kikumi Tatusoka who came to ETS to join Charlie Lewis, her former dissertation advisor at the University of Illinois. DY learned many things from Bob Mislevy and they have been leading an annual NCME training session based on their book Bayesian Networks in Educational Assessment since 2002 (Almond et al. 2015). Charlie introduced DY to the world of CAT and they developed the tree-based CAT algorithm. She was always impressed by how Bob and Charlie solved problems. She was also impressed about 20 years ago, when Charlie hosted his former dissertation advisor John Tukey (from Princeton University in 1970) at ETS once a month to consult on their projects. DY brought modern computer outputs with analyses results and plots to show John. John didn't look at those outputs, instead he took a piece of paper and a pencil then started to draw a stem-leaf graph, and he asked everyone what they thought the results should be, which were the results DY produced after hours of computing! DY was astonished by how he explained things from his head, which is the way Charlie writes out the equations from his head at any point! She wanted to learn more. So, she later followed Charlie to Fordham University to finish her PhD in Psychometrics with her dissertation on computerized multistage testing (MST) which was co-advised by Charlie Lewis and Alina von Davier. They subsequently published a book (Yan et al. 2014) and DY received the 2016 AERA Significant Contribution to Educational Measurement and Research Methodology Award. DY was also honored to receive 2011 ETS Presidential Award, 2013 NCME Brenda Loyd Dissertations Award, and 2015 IACAT Early Career Award. Currently, she is responsible for ETS's automated scoring systems evaluations and analyses including a *Handbook of Automated Scoring: Theory into Practice* (Yan et al. 2020).

During her career, DY faced many challenges such as work and life balance including operational work versus research and development, schedule conflicts, family, and child raising. From her work on many operational programs and research and development projects, she gained experiences dealing with real-world practical issues and finding solutions. These helped her to create more innovative research questions and to develop and implement systems that increase accuracy and efficiency by using optimization and automation. Her daughter Victoria Song often slept on her desk or on the office floor. She grew up at ETS, volunteered

and interned at ETS, and is working on her dissertation advised by Fordham Anne Anastasi Chair Professor David Budescu. All DY's learnings and experiences are good lessons in her life. She appreciates her mentors who had great influence in her career and life.

5 Advice, Recommendations, and Lessons Learned

Although we all have worked at different academic departments and in different countries, we still have similar experiences and we have learned many things during our journeys. The some of the lessons all seven of us have learned are described below.

5.1 Things to Do

Dare to say yes to exciting projects and decline administrative committees if they just need a woman and not specifically your competence.

Try to find a supportive work environment with people you can be on the same level with. When you are young try to find a good mentor and once you are older try to be a good mentor: keep an open mind and learn. You never know when they become fruits in your life.

Work with people you enjoy spending time with and those you dare to say that you do not understand what they mean. It is more fun and it is more rewarding for both partners.

Don't be afraid to collaborate with new people. Some of the best collaborations come from just listening to a conference presentation and suggesting to collaborate with a joint topic, even if none of the people knew each other before.

Probably the most important lesson to share with junior psychometricians is to believe, respect, and acknowledge one's own ideas and at least in so much as to try them out. This would start by just writing down the idea, and then write the computer code, prove the theorem, and/or test the result empirically.

The "just do it!" approach is usually good. Even if an idea is not valuable but by trying them out one builds both expertise and confidence. It is tempting for a novice to talk about an idea but not pursue it, even in the face of statements in the literature that it is not possible, not feasible, or not true.

We all believe that for an academic career, what you study is your choice and therefore it is important to choose what interests you. Always go back to original sources and thoroughly read the literature. Do not rely solely on search engines because you might miss connections between different literatures.

Balancing family and career is possible, challenging, and rewarding. Finding a balance that works can be the harder part and this will change over time. A supportive husband or partner, friends, mentors, and colleagues, as well as a flexible

work schedule are invaluable. When looking for a position, consider the attitudes toward working women and family policies at the institution and laws within that country. For example, several of us choose to be parents and to be professors. The choice to be a parent can impact research productivity but recognize that this is temporary and does not imply that you are not serious about your career. It is not strange that there might be gaps in productivity, which coincide with major life events. Life does not always conform to an academic calendar. But planning can help your career a lot even when life events happen.

5.2 Things Not to Do

During our careers, there are also things we have learned that is better to avoid, and below is a short list of some "don'ts":

Don't say yes to committees just because they need to fill the female or minority spot, unless you really want to do the work.

Don't say no to something you wish to do because you have never done it before or you are unsure about your capacity. If you are interested in the topic you will learn during the process.

Don't let shyness or modesties stand in your way of your achievements. Many of us may be introverts and find it uncomfortable to present our work in public; however, recognize that those who are in the audience want to hear about your accomplishments.

Don't just hang with the crowd you know at conferences. Try to meet and engage with new people in the area which interest you.

Don't only attend sessions in your own area. Attending other sessions is an opportunity to learn and expand your knowledge.

Don't be selfish in collaborations, especially with younger researchers. Your generosity most likely will be rewarded later.

Don't discount or underestimate your knowledge. If you are in a meeting you are probably there as an expert.

Don't always believe what you read in the literature. Knowledge and understanding evolve over time, and mistakes do sometimes slip by reviewers and editors.

Don't despair if you get a reject/revise on a submitted paper. This means a bit more work and it will probably get published.

6 Future Directions

In the future, it is important to help our peers: to support young researchers and to help to build organizational structures to promote a healthy career. As senior researchers we should be aware of gender inequalities and make sure that scientific

conference program organizers make significant efforts to represent both genders and their scientific contributions in the keynote and invited talks. This goal serves to promote and acknowledge the work done by women researchers but also very importantly for creating role models for the younger generations. On one hand, we would like to acknowledge that we are fortunate to see how the science environment is changing and becoming more inclusive, while preserving and applying the high standards to all. On the other hand, it is crucial to continue to address the issues of gender and other inequalities that characterize most aspects of our jobs including recruitment, promotions, opportunities for collaborations, publishing our work, and other contributions to our respective working environments. Part of it is to understand that gender inequality has a negative impact on our profession and society. If we believe that our fields are exciting and important and that impact our and future generations then we should get all the talent we can get. There are many historical reasons as discussed in the introduction for the gender gap. We need to continue addressing how important it is for our generation of women to take an active role in promoting women's work and contributions, for mentoring women to help them progress and get promoted. Achieving these goals that are within our means can create a more balanced and healthy working environment and society for all.

There are also many systematic initiatives from recognized professional bodies. The London Mathematical Society (LMS) is committed to actively addressing the issues facing women in mathematics. It is concerned about the loss of women from mathematics, particularly at the higher levels of research and teaching, and at the disadvantages and missed opportunities that this represents for the advancement of mathematics. The LMS Council Statement on Women in Mathematics recognizes the need to give active consideration to ensuring that men and women are treated equally in their prospects, recognition, and progression.

The Association for Women in Mathematics' purpose (1971) is to encourage women and girls to study and pursue careers in the mathematical sciences, and to promote equal opportunity and equal treatment of women and girls in the mathematical sciences. There is also the "This is Statistics" campaign to pitch Big Data professions to middle and high school girls and minorities. This is very important since Data Science and Big Data analysis is an emerging field. Other initiatives include a yearly conference: Women in Statistics and Data Science (since 2016). The R-Ladies is a worldwide organization whose mission is to promote gender diversity in the R community.

Among the things we do and we should continue doing: address stereotyping in educational and training choices at school (and at home) at a young age, adopt teaching strategies to increase engagement of girls in mathematics, act as role models, achieve a better gender balance of teaching at all levels of education, and promote STEM professions among young women. In addition, we should organize and run regular workshops at conferences with themes that provide training on leadership skills (how to be influential and impactful): career events and workshops focusing on female students and junior academics on how to empower women and minorities. The importance of role models: strong representation of women

in keynote and invited talks as well as larger representation of women in editorial boards and editorships.

It is not enough to increase the quotas for female participation. We also need to create an environment in which women will have an equal voice and can prosper in their careers and personal lives, which is linked to the rate and time of promotions for women. To quote the character from Ratatouille, "anyone can be a chef, but not everybody can be a chef."

References

Almond, R., Mislevy, R., Steinberg, L., Yan, D., & Williamson, D. (2015). *Bayesian networks in educational assessment*. New York: Springer.

Anderson, C. J., & Vermunt, J. K. (2000). Log-multiplicative association models as latent variable models for nominal and/or ordinal data. *Sociological Methodology, 30*, 81–121.

Black, C., & Islam, A. (2014). Women in academia: What does it take to reach the top? *The Guardian*. Retrieved from https://www.theguardian.com/higher-education-network/blog/2014/feb/24/women-academia-promotion-cambridge

Cochran, W.G., & Cox, G.M. (1957). Experimental Design. 2nd Edition, John Wiley and Sons, New York.

González, J., & Wiberg, M. (2017). *Applying test equating methods – Using R*. Cham: Springer.

Holzinger, K. J., & Swineford, F. (1937). The bifactor method. *Psychometrika, 2*, 41–54. https://doi.org/10.1007/BF02287965.

Huerta-Sanchez, E., & Rolfs, R. (2019). *Hidden contributor to professor: An interview with Margaret Wu. From genes to genomes*. Retrieved from http://genestogenomes.org/margaret-wu

Mason, M. A., & Wolfinger, N. H. (2013). *Do babies matter? Gender and family in the Ivory tower*. New Brunswick: Ruthers University Press.

McCurry, J. (2019). Women outperform men after Japan medical school stops rigging exam scores. *The Guardian*. Retrieved from https://www.theguardian.com/world/2019/jun/19/women-outperform-men-after-japan-medical-school-stops-rigging-exam-scores

Okahana, H., & Zhou, E. (2017). *Graduate enrollment and degrees: 2006 to 2016*. Washington, DC: Council of Graduate Schools.

Rubini, M., & Menegatti, M. (2014). Hindering women's careers in academia: Gender linguistic bias in personnel selection. *Journal of Language and Social Psychology, 33*(6), 632–650.

van den Brink, M., Benschop, & Janssen, W. (2010). Transparency as a tool for gender equality. *Organization Studies, 31*, 1459–1483.

Yan, D., von Davier, A., & Lewis, C. (Eds.). (2014). *Computerized multistage testing: Theory and applications*. Boca Raton: Chapman and Hall.

Yan, D., Rupp, A. A., & Foltz, P. (2020). *Handbook of automated scoring: Theory into practice*. *Assessment, 23*, 279–291. Taylor & Francis. https://doi.org/10.1177/1073191115583714.

Yong, E. (2019). The women who contributed to science but were buried in the footnotes. *The Atlantic*. Retrieved from https://www.theatlantic.com/science/archive/2019/02/womens-history-in-science-hidden-footnotes/582472

Developing a Concept Map for Rasch Measurement Theory

George Engelhard Jr and Jue Wang

Abstract The purpose of this paper is to identify and describe the key concepts of Rasch measurement theory (Rasch G, Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research, Copenhagen. (Expanded edition, Chicago: University of Chicago Press, 1980), 1960/1980). There have been several taxonomies describing item response theory (Kim S-H et al., A taxonomy of item response models in Psychometrika. In: Wiberg M, Culpepper S, Janssen R, Gonzáles J, Molenaar D (eds) Quantitative psychology: 83rd annual meeting of the Psychometric Society. Springer, New York City, pp 13–23, 2019; Thissen D, Steinberg L, Psychometrika 51:567–577, 1986; Wright BD, Masters GN, Rating scale analysis: Rasch measurement. MESA Press, Chicago, 1982), and this paper extends these ideas with a specific focus on Rasch measurement theory. Rasch's measurement work reflects a key milestone in a paradigmatic shift from classical test theory to item response theory (van der Linden WJ, Handbook of item response theory, volume 1: models. CRC Press, Boca Raton, 2016). We include a categorization of measurement models that are commonly viewed as Rasch models (dichotomous, rating scale, partial credit, and many-faceted), as well as extensions of these models (mixed, multilevel, multidimensional, and explanatory models). Georg Rasch proposed a set of principles related to objectivity and invariance that reflect foundational concepts underlying science. Rasch measurement theory is the application of these foundational concepts to measurement. Concept maps provide useful didactic tools for understanding progress in measurement theory in the human sciences, and also for appreciating Rasch's contributions to current theory and practice in psychometrics.

G. Engelhard Jr (✉)
Quantitative Methodology Program, Educational Psychology, The University of Georgia, Athens, GA, USA
e-mail: gengelh@uga.edu

J. Wang
Research, Measurement, and Evaluation Program, The University of Miami, Coral Gables, FL, USA
e-mail: jue.wang@miami.edu

M. Wiberg et al. (eds.), *Quantitative Psychology*, Springer Proceedings in Mathematics & Statistics 322, https://doi.org/10.1007/978-3-030-43469-4_2

Keywords Rasch measurement theory · Philosophy of measurement · Invariant measurement

1 Introduction

> *The concept of "objectivity" raises fundamental problems in all sciences. For a statement to be scientific, "objectivity" is required.* (Rasch 1964, p. 1)

Rasch described several models for measurement that he developed to address problems encountered in his research work. His seminal book entitled *Probabilistic Models for Some Intelligence and Attainment Tests* (Rasch 1960/1980) introduced several models of measurement including models for misreadings, reading speed, and item analysis. These models became the basis for numerous advances in measurement theory.

Rasch measurement theory has been described as "a truly new approach to psychometric problems ... [that yields] non-arbitrary measures" (Loevinger 1965, p. 151). As pointed out by van der Linden (2016), the first chapter of Rasch's book is required reading for anyone seeking to understand the transition from classical test theory to item response theory (IRT). In his words, "One of the best introductions to this change of paradigm is Rasch (1960/1980, Chapter 1), which is mandatory reading for anyone with an interest in the subject" (van der Linden 2016, p. xvii). Wright (1980) commented that Rasch's psychometric methods "go far beyond measurement in education or psychology. They embody the essential principles of measurement itself, the principles on which objectivity and reproducibility, indeed all scientific knowledge, are based" (p. xix). This study explores what Rasch did to receive these accolades.

In order to explore current perspectives on Rasch measurement theory, we conducted a Web of Science search using the topic phrase "Rasch measurement theory". This bibliometric search was limited to the twenty-first century (2000–2019), and 754 references were identified. Figure 1 shows frequency of articles related to Rasch measurement theory. It is also interesting to note the distribution of these articles over various fields with psychology (N = 240), health care sciences (N = 125), and educational research (N = 109) identified as the top three areas.

The purpose of this study is to identify the key concepts that define Rasch measurement theory. Specifically, the following questions guide our research: (a) What is Rasch measurement theory? (b) What are the key concepts that define Rasch measurement theory?

2 What Is Rasch Measurement Theory?

One way to define Rasch measurement theory is by the specific models for measurement proposed by Rasch, and also the models that are considered extensions

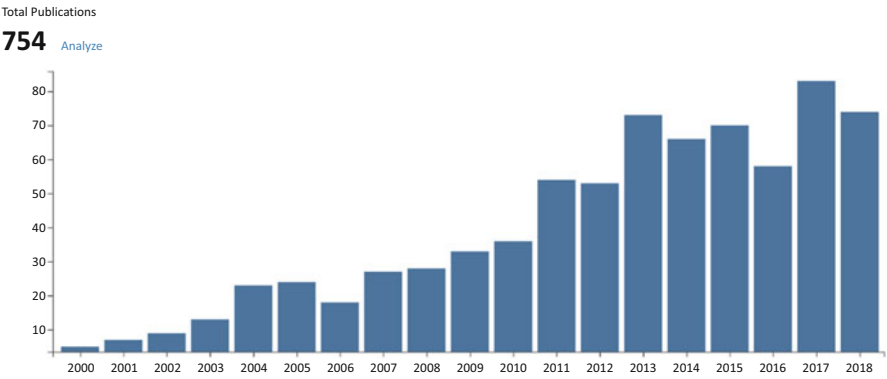

Fig. 1 Frequency of articles on Rasch measurement theory (Web of Science)

Name	Log-odd Forms of Rasch Models
Dichotomous Model	$Ln\left(\frac{P_{ni1}}{P_{ni0}}\right) = \theta_n - \delta_i$
Partial Credit Model	$Ln\left(\frac{P_{nik}}{P_{nik-1}}\right) = \theta_n - \delta_{ik}$
Rating Scale Model	$Ln\left(\frac{P_{nik}}{P_{nik-1}}\right) = \theta_n - (\delta_i + \tau_k) = \theta_n - \delta_i - \tau_k$
Many Facet Models	
Many Facet Partial Credit Model	$Ln\left(\frac{P_{nmik}}{P_{nmik-1}}\right) = \theta_n - \lambda_m - \delta_{ik}$
Many Facet Rating Scale Model	$Ln\left(\frac{P_{nmik}}{P_{nmik-1}}\right) = \theta_n - \lambda_m - (\delta_i + \tau_k) = \theta_n - \lambda_m - \delta_i - \tau_k$

Fig. 2 Commonly used Rasch Models. θ_n= person ability measure; δ_i= difficulty of item i; δ_{ik}= difficulty of step k of item i (assuming unique scale structure of each item); τ_k= difficulty of step k (assuming common scale structure among all items); λ_m= scoring severity of rater m

of the unidimensional Rasch model. One of the earliest taxonomies of Rasch models is offered by Wright and Masters (1982). They described a family of Rasch models designed to analyze dichotomous and polytomous responses obtained from persons based on items that are developed to represent a unidimensional continuum. Specifically, Wright and Masters (1982) described five Rasch models: Dichotomous, Partial Credit, Rating Scale, Binomial Trials, and Poisson Count. Linacre (1989) extended this family of Rasch models to include raters. Figure 2 describes the most commonly used Rasch models.

Kim and his colleagues (2019) categorized IRT articles appearing in *Psychometrika* from 1960s to 2010s. They identified 157 articles related to Rasch measurement theory. About 41.64% of the total IRT articles in *Psychometrika* are related to Rasch measurement theory. Figure 3 shows the frequency of articles on Rasch measurement theory over time by model type.

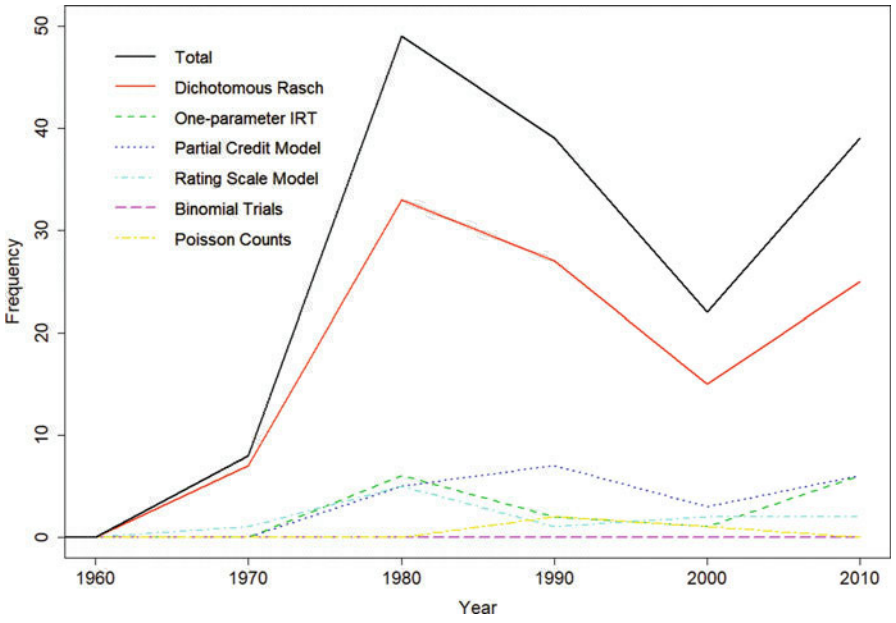

Fig. 3 Frequency of articles on Rasch models over time published in *Psychometrika* (Kim et al. 2019). Total (black solid line) shows the frequency of all Rasch models over time. Further bibliometric evidence shows that Rasch measurement theory continues to influence measurement research as provided by Aryadoust and Tan (2019)

There have been numerous extensions to Rasch models. Here is a partial list of the extensions: (a) mixed Rasch model (Rost 1990), (b) multilevel Rasch measurement model (Adams et al. 1997b), and (c) multidimensional random coefficients multinomial logit models (Adams et al. 1997a). Since research continues on extensions to Rasch measurement theory, this list should be considered incomplete.

In addition to defining Rasch measurement theory based on models for measurement that specifically include Rasch's name, a complementary approach is to consider the key concepts that define Rasch measurement theory. These key concepts considered in the next section are based on Rasch's views of science and the application of these concepts to measurement.

3 What Are the Key Concepts that Define Rasch Measurement Theory?

Looking then for concepts [of measurement] that could possibly be taken as primary it seems worthwhile to concentrate upon two essential characteristics of "scientific statements" 1. they are concerned with "comparisons"; 2. the statements are claimed to be "objective"; both terms of course calling for precise qualifications. (Rasch 1964, p.2)