

Institution: University of Kent		
Unit of Assessment: 4 – Psychology, Psychiatry and Neuroscience		
Title of case study: Improving Personality Assessment Tools for World-Leading Test Publishers		
Period when the underpinning research was undertaken: 01/05/2012 to date		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Dr Anna Brown	Reader in Psychological Methods and Statistics	01/05/2012 to date
Period when the claimed impact occurred: 01/08/2013 to 31/12/2020		
Is this case study continued from a case study submitted in 2014? No		
1. Summary of the impact (indicative maximum 100 words)		
<p>Research by Dr Anna Brown led to the development of novel personality assessment tools that are resistant to response biases and 'faking good' by candidates. Brown's methodology enabled multiple world-leading test publishers and organisations, including SHL and Korn Ferry, to develop solutions that improve the validity of proprietary tests. Over five million individual psychological assessments incorporating this methodology have taken place in 37 different languages across 40 different countries since August 2013. Publishers commend Brown's 'inspiring insights and actionable recommendations' [b], in recognition of her capacity to help them design 'wonderful, innovative' solutions [d] that maximise validity and fairness in assessment.</p>		
2. Underpinning research (indicative maximum 500 words)		
<p>Dr Anna Brown's research directly addresses the scientific and practical challenges of psychological assessment. She is a psychometrician who focuses on modelling response processes to non-cognitive assessments (e.g. personality, motivation, attitude) using multidimensional item response theory.</p> <p><i>Ipsative</i> (or relative-to-self) questionnaires ask respondents to compare sets of two or more stimuli from the same domain, such as behaviours, values, or interests. For example, to measure occupational interests, respondents may be asked to indicate their preference for two or more activities such as: (a) planting roses; (b) receiving telephone calls; or (c) building bridges. Preferences can be expressed as rank orders (e.g. A>C>B), or graded in terms of strength (e.g. 'prefer a little' A to C, or 'prefer a lot' C to B), or as a percentage of the total (e.g. 50-40-10). These <i>forced-choice</i>, <i>graded preference</i>, and <i>compositional</i> formats, respectively, have been popular with practitioners since the 1940s because they elicit more differentiated judgements (Kahneman, 2011) and prevent many response biases. However, they have also been the focus of longstanding controversy and criticism from the academic community, which has argued that while ipsative questionnaires are good for measuring <i>intra</i>-individual differences (different attributes within the same person), they are severely limited for measuring <i>inter</i>-individual differences (the same attribute between different people; e.g. Closs, 1966; Johnson, Wood, & Blinkhorn, 1988; Meade, 2004). In contrast, Brown's methodology enables proper scaling of forced-choice data and allows ipsative questionnaires to be used for inter-individual comparisons.</p> <p>Brown's background in the personnel assessment sector, ongoing programme of consultancies, and Scientific Advisory Board memberships with test publishers ensure a close working relationship with the organisations that use her research. Brown joined the University of Kent from a senior post in industry in 2012 and has since developed and published a unified theory [R6] that can be applied to any type of ipsative data, based on utility judgement and discriminant process proposed by Thurstone (1927).</p>		

Her research at Kent resulted in the discovery of fundamental rules for identifying the scale origin in comparative judgements [R6] and for computing test information and reliability [R1]. Two important instances of the unified approach were addressed by the development of methods for scaling *compositional* formats [R5], and for scaling *graded preferences* [R1]. Further applied research has informed the use of comparative judgements to increase validity of computer-adaptive personality testing [R4], and organisational appraisals [R3]. These important advances in comparative measurement methodology and practical know-how to develop effective forced-choice tests [R2] form the basis for this case study.

Brown has developed these novel insights through formal partnerships with test publishers and organisations that use psychological testing, which have led to enhancements in the design of their proprietary assessment tools (as detailed below). The key underpinning research has also been documented through academic publication in prestigious journals or handbooks.

3. References to the research (indicative maximum of six references)

[R1] Brown, A., & Maydeu-Olivares, A. (2018). Ordinal Factor Analysis of Graded-Preference Questionnaire Data. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 516-529. doi: 10.1080/10705511.2017.1392247

[Ordinal Factor Analysis of Graded-Preference Questionnaire Data - Kent Academic Repository](#)

[R2] Brown, A., & Maydeu-Olivares, A. (2018). Modeling forced-choice response formats. In Irwing, P., Booth, T., & Hughes, D. (eds.), *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*, pp. 523-570. London: John Wiley & Sons.

[Modeling forced-choice response formats - Kent Academic Repository](#)

[R3] Brown, A., Inceoglu, I., & Lin, Y. (2016). Preventing Rater Biases in 360-Degree Feedback by Forcing Choice. *Organizational Research Methods*, 20(1), 121-148. doi: 10.1177/1094428116668036

[Preventing Rater Biases in 360-Degree Feedback by Forcing Choice - Kent Academic Repository](#)

[R4] Lin, Y., & Brown, A. (2016). Influence of Context on Item Parameters in Forced-Choice Personality Assessments. *Educational and Psychological Measurement*, 77(3), 389-414. doi: 10.1177/0013164416646162

[Influence of Context on Item Parameters in Forced-Choice Personality Assessments - Kent Academic Repository](#)

[R5] Brown, A. (2016). Thurstonian Scaling of Compositional Questionnaire Data. *Multivariate Behavioral Research*, 51(2-3), 345-356. DOI: 10.1080/00273171.2016.1150152

[Thurstonian Scaling of Compositional Questionnaire Data - Kent Academic Repository](#)

[R6] Brown, A. (2016). Item Response Models for Forced-Choice Questionnaires: A Common Framework. *Psychometrika*, 81(1), 135-160. doi: 10.1007/s11336-014-9434-9

[Item Response Models for Forced-Choice Questionnaires: A Common Framework - Kent Academic Repository](#)

Grants and Awards

[G1] Dr Anna Brown has carried out a series of consultancies for organisations included in this case study between August 2013 and December 2020, offering methodological advice. Her clients include Caliper, the German Federal Employment Agency, the Enrolment Management Association, ECHO Listening, Talent Q, Whelan & Associates, and Talentoday.

4. Details of the impact (indicative maximum 750 words)

Since **2014**, several world-leading test publishers, including SHL and Korn Ferry, have applied Dr Anna Brown's research to enhance or develop their assessment tools. The adoption of Dr Brown's findings by these international stakeholders means that her methodologies are now in use across 40 countries and in contexts including recruitment, educational selection, and personal development, with over one million of these tests administered each year [a].

The specific benefits of Brown's research vary between test publishers, depending on their mission and whether they market proprietary psychological assessment products. On this basis, in some instances much of her method is incorporated into specific tests as the result of consultancies, whilst in others her methods are adopted by accessing her research through the scientific literature. Typically, her methodology is used to produce assessment tools that capitalise on the advantages of comparative judgements, such as robustness to response biases, and with defensible measurement properties and validity.

Below we demonstrate some notable outcomes for the organisations that use Dr Brown's research.

SHL is one of the world's largest occupational test publishers, operating across Europe, North America, the Middle East, South Africa, and Asia Pacific. Since **2016**, Dr Brown has provided 'inspiring insights and actionable recommendations', to help SHL refine its OPQ32r assessment based on Thurstonian scaling methodology, with over one million individual candidate tests administered annually (i.e. an estimated 5 million in total) [b]. Dr Brown also worked with SHL to develop the theory behind computerised adaptive testing with comparative judgements [R4] to move away from fixed to adaptive assessments (patented in the United States, number 10460617). More recently, in response to the COVID-19 pandemic, SHL applied Brown's scaling method to launch the RemoteWorkQ – a free-access questionnaire designed to identify an individual's strengths and areas of development in a remote working environment, to help people work more productively from home. As SHL's Managing Research Scientist states: 'The multidimensional IRT model for forced-choice response data developed by Anna [Brown] and Prof. Alberto Maydeu-Olivares is truly revolutionary. I have applied the model in my work extensively, and am impressed by the robustness and flexibility it offers.' [b]

Korn Ferry is also one of the world's largest occupational test publishers, operating across the US, Latin America, the Middle East, and Asia Pacific. In **2014**, the company developed two new proprietary tests taking advantage of the scaling method developed by Brown [R6, c]. The tools were made available to the executive search market in **2015** and the broader talent management market in **late 2016/early 2017**. Since then, Korn Ferry has administered 250,000 assessments based on this new methodology [c].

Caliper is an employee assessment specialist with a 60-year track record in the field and is the publisher of the industry leading 'Caliper Profile' test [f]. A statement from Caliper details the impact of a series of consultancies that Dr Brown carried out between **August 2013** and **July 2017**:

'Anna was able to use Thurstonian IRT to much more accurately estimate the psychometric properties of our items and scales [...]. We continue to this day employing the recommendations made and Anna continues to advise us along the way. The redesign and changes to the format of the questionnaire has been ongoing, but a major benefit based on Anna's recommendations is a reduction in the time to complete the assessment as well as a reduction in the cognitive complexity to the test taker, which leads to an overall better candidate experience. Furthermore, we were able to leverage Anna's thorough analyses to provide more accurate information about the scale properties in our most recent technical manual. Anna's methodology and consultancy has helped us in our strategic priority in developing an assessment that will serve to reinforce our strengths, address our weaknesses, strengthen our current competitive advantage, and allow us to continue to be a major player in the end-to-end talent management and consulting market segment' [f].

Talentoday is a French predictive analytics start-up consultancy with a proprietary instrument that has been challenged by some large clients because of its ipsative structure. A collaboration with Brown in **2017-18** led to the development of MyPrint, which eliminated the problems of ipsative data. As the Chief Scientist at Talentoday stated: 'Dr. Anna Brown then came up with a wonderful, innovative solution: she proposed to use a multidimensional 2-forced-choice format with *graded preferences*' [R1, d]. In **2018**, **Medix** (a US workforce solutions and staffing company that works in the healthcare, science, and IT sectors) acquired Talentoday in order to add this highly innovative instrument to its portfolio [d].

The Enrolment Management Association (EMA) develops assessment tools for selection into independent schools across the US. Brown joined the Technical Advisory Panel of EMA in **2018** to help develop a personality assessment tool called 'Character Skills Snapshot'. The Senior Research Scientist at EMA states that Brown's 'methodology has guided how we write our statements, how many statements to write, and the claims we make about the validity of scores given the high-stakes admissions context' [e].

ECHO Listening is a US management training specialist that publishes the communication development tool, the 'ECHO Listening Profile', which incorporated Dr Brown's ipsative model in **2018** to improve validity and to keep questionnaires shorter, which the company is 'finding to be a strong value-add for those wanting to deploy ECHO in the business world' [h].

In addition to the detailed examples indicated here, a substantial number of other organisations have also applied and benefited from Brown's methodologies. These include:

Whelan & Associates – a US consultancy specialising in building selection, retention, and employee engagement tools [g];

The German Federal Employment Agency (Psychological Research and Development department responsible for skills and competence assessment of job seekers) [i];

Talent Q – a personnel assessment specialist providing services to the largest public and private organisations in the Russian Federation and neighbouring states [j].

5. Sources to corroborate the impact (indicative maximum of 10 references)

[a] Summary table of test publishers and assessments referred to in this case study.

[b] SHL: statement from the Managing Research Scientist describing the impact of Brown's research on the company's proprietary personality tests.

[c] Korn Ferry: *Assessment of Leadership Potential: Research Guide and Technical Manual* (2015) – [R6] referenced on p. 33 (but incorrectly dated 2014, not 2016); and statement from the Director Talent Solution Design describing the impact of Brown's research on the company's proprietary personality tests.

[d] Talentoday: *My Print Short Technical Manual* (2018) – [R1] referenced on p. 30 (but dated 2017, not 2018); and statement from the Chief Scientist describing the impact of Brown's research on the company's proprietary personality test.

[e] Enrolment Management Association: statement from the Chief Testing and Research Officer describing the impact of Brown's research on the company's proprietary assessment tool.

[f] Caliper: statement from Senior Consultant (Psychometrics & Data Methods) and Senior Research Scientist describing the impact of Brown's research on the company's proprietary personality test.

Impact case study (REF3)

[g] Whelan & Associates: statement from the President describing the impact of Brown's research on the company's proprietary personality test.

[h] ECHO Listening: statement from the Chief Operating Officer describing the impact of Brown's research on the company's proprietary assessment tool.

[i] German Federal Employment Agency: statement from the Scientific Officer on the impact of Brown's research on the organisation's assessment tool.

[j] Talent Q: statement from the General Manager describing the impact of Brown's research on the company's proprietary personality test.