



The Response Process in Psychometric (Intelligence) Testing and Validity



Vanessa Torres van Grinsven^{1,2*}

¹Faculty of Psychology Open University, Netherlands

²Department of Special Education and Rehabilitation, University of Cologne, Germany

Submission: May 15, 2023; Published: June 01, 2023

*Corresponding author: Vanessa Torres van Grinsven, University of Cologne, Faculty of Human Sciences, Department of Special Education and Rehabilitation, Section Research Methods, Aachenerstraße 197-199 20G, 50933, Cologne, Germany, Email: v.torresvangrinsven@uni-koeln.de

Abstract

Standardized (psychometric) tests are common tools used in the diagnosis of learning and behavior disabilities in children. In this essay, I argue that, in a test-taking situation with a psychometric test, complex interrelationships take place. This complex interaction shapes the response process which results in a performance, i.e., the results of this test. Current psychometric practices and test-taking practices and interpretation of results fail to take this response process sufficiently into account. I present the interactionist and process-performance approach as a framework to help address this phenomenon, discuss a few examples of subtests of WISC IV based on this approach, and call for more research into the response process and aspects of tests that may impact validity.

Keywords: Diagnostic test; Psychometric test; Standardized test; Response process; Social interaction; Performance; Validity; WISC

Introduction

When diagnosing and treating learning and behavior disabilities, psychometric tests are commonly used diagnostic tools. Research has however shown that important errors may occur despite the application of validation processes and adherence to quality criteria for psychometric tests [1]. In a seminal work about the construction of questions for (standardized) questionnaires used in social and behavioral research, Foddy [2] presents the basic assumption in the science of questionnaire construction that question-answer behavior involves complex interrelationships between sociological, psychological and linguistic variables [2], which (my addition) results in a response process.

I argue that, similarly, in a test-taking situation such complex interrelationships take place. This means that, in such a test-taking situation, a complex interaction takes place between the test-taker and the test. And, in the case of most tests used with children, there's also an assessor involved. This complex interaction then shapes the (response) process that shapes the performance and hereby the results of the test [1]. Here, a lot of factors may come into play, such as personality traits, motivational and emotional variables, and – related to all these – sociocultural traits.

It is not only about the characteristics of test-items themselves, or the way properties of the test-items interact with the assessors and test-takers, or the relationship between the assessors and the test-takers. All these elements constitute a dynamic, interrelated set of elements.

At the same time, it seems that response processes and test consequences have been largely neglected in validation practice in the area of psychology [3,4]. Nonetheless, in the use of psychometric tests for diagnostics, and especially with children with suspected intellectual and developmental disabilities, adopting the process-performance approach [1] and an interactionist viewpoint, will give great chances for improvement of diagnostics and hereby prevention and treatment of intellectual and developmental disabilities.

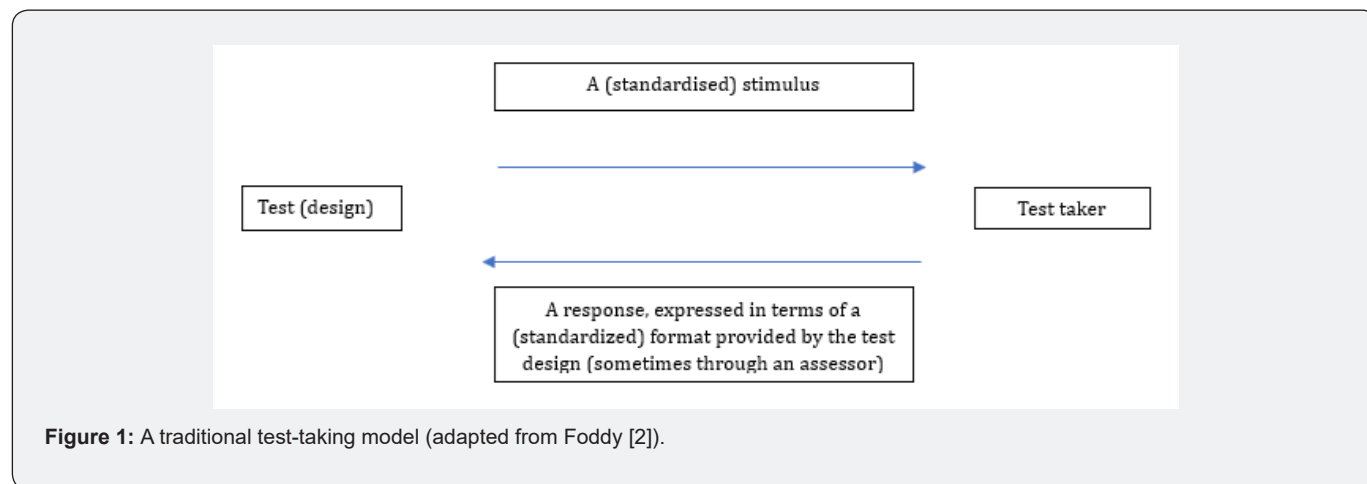
Interactionism and the process-performance approach

The term 'symbolic interactionism' was initially described by sociologist Herbert Blumer [5,6], who endorsed a set of ideas first put forward by the social philosopher George Herbert Mead. An important ground of the symbolic interactionist approach is

that human beings interpret and define each other's actions. They do not merely simply react to each other in a simple stimulus – response way, and therefore responses are not made to acts but rather to interpreted acts. Importantly, these processes occur in all social situations, including a test-taking situation and thus test-taking behavior. Consequently, these processes form the test-taking situation as a social interaction, or, better said, a

sociocultural interaction.

A traditional test-taking model that implicitly seems to underline standardized psychometric testing can be expressed like this model resembling the traditional survey model from Foddy [2] (Figure 1).



However, in this view, the underlying response process has been neglected. From an interactionist viewpoint the response process may become especially complex in the case of psychometric testing of children, i.e., when the test taker is a child being assessed. Especially in the case of children suspected of having a developmental or intellectual disability, it would be important to acknowledge this response process. One reason is the presence of an assessor. Another one would be that taking a test in different forms requires diverse sets of skills. These diverse sets of skills that are required for different tests may sometimes be unrelated or lay outside the dimension that is formally being tested. A child with a developmental or intellectual disability may have more chances of lacking these skills and therefore scores of a test may be invalid.

Let's look at intelligence tests for example. Psychometric intelligence tests are commonly used measures. The Wechsler Intelligence Scale for Children, and the Woodcock-Johnson Tests of Cognitive abilities are the most commonly used intelligence tests for children [7]. Considering intelligence, personality and intelligence have traditionally been conceptualized as distinct domains (though also critically discussed by [8,9]), and the typical measurement of intelligence through tests is supposed to reflect 'maximal performance' [8]. This means performance when individuals are trying their hardest.

However, the interactionist and process-performance approach [1,10,11] point to a situation where the performance in a test is shaped by the response process as a social interaction and thus the results of a test might or might not indicate this maximal

performance. Research shows that several factors that might come into play here are personality traits, motivational and emotional variables, but – and related to these – also socio-cultural factors (see [1,9] for a review of several of those factors).

Content versus Form

Let's have a look at the content and form (or format) of a few subtests of WISC IV [12] to see an example of how personality traits and socio-cultural factors might come into play, in this interaction and response process.

The WISC-IV is composed of indexes, that are composed of subtests composed of items. Each one of the indexes and subtests taps for different dimensions (content) and each one has its own way of 'questioning' (form or format). For example, for the VCI (Verbal Comprehension Index) children must answer orally presented questions (form) that assess among others common-sense reasoning (content). To perform these tasks (the form) certain abilities and skills are required that might fully or only to a certain extent, or not at all, be part of the dimension being assessed.

The PRI (Perceptual Reasoning Index) is another index of the WISC IV. Its primary purpose is to examine nonverbal so-called fluid reasoning skills (content). The index also requires visual perception and organization and reasoning with visually presented, nonverbal material to solve problems. One of its subtests, the Block Design subtest, also requires visual-motor coordination and the ability to apply all skills in a quick, efficient manner.

There is some evidence that the Matrix Reasoning (MR) and Pictures Concepts, that are also subtests of the PRI, measure verbal ability to a degree, as these tasks may involve subvocal verbal reasoning [13-15]. Thus, these measures may be impacted by verbal mediation and verbal skills. That is, the format of the tests influences the assessment of the dimension tapped (content). Other research shows that time restrictions (format) in an intelligence test might interact with personality traits (see e.g. [1,9]) which influences the results.

When looking at the format of this MR subtest in detail, more questions may arise concerning the format. First, the Matrix Reasoning subtest is a kind of multiple-option single-choice test. Kubinger & Wolfbauer [16] show that some interaction may occur between personality traits and instructions and results of an adapted Culture Fair Test of Cattell [17], a test that has a multiple-option single choice response format.

What's more, despite the PRI to be posited to tap the kind of problem-solving not taught in schools [18], culture-based socialization might also impact how one handles the response option in such tests [16] and this might also be true for the Matrix Reasoning subtest. Moreover, solving the MR subtest, requires a perception of linear order. However, from a cultural anthropologist's viewpoint linear thinking is culture-based. Also, the MR subtests are similar to making a puzzle. It can be similarly argued that making puzzles is a socio-cultural activity and thus a socio-cultural based skill. I.e., Raven's matrices may be considered cultural artifacts [14] and so the MR subtest [19-21].

Conclusion

We need more research into the response process occurring in psychometric testing and especially into validity issues that are related to this response process. The interactionist and process-performance approach are helpful frameworks as they allow us to see the test-taking situation as a complex social (sociocultural) interaction that shapes the response process and hereby the results of a test.

References

- Torres van Grinsven V (2022) Sources of measurement error in pediatric intelligence testing. *Methodological Innovations* 15(1): 96-104.
- Foddy, WH (1994) *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research*. New Edition, Cambridge University Press, Cambridge, UK.
- Hubley AM (2018) Missed opportunities in testing and assessment: Response processes and test consequences. Keynote address at the International Congress on Applied Psychology (ICAP), Montréal, Canada, PQ.
- Zumbo BD, Chan EKH (2014) *Social indicators research series: Vol. 54. Validity and validation in social, behavioral, and health sciences*. Springer International Publishing.
- Blumer H (1967) *Society as Symbolic Interaction*. In: J Manis, B Meltzer (Eds.), *Symbolic Interaction: A Reader in Social Psychology*, Boston; Allyn and Bacon, (1st edn).
- Blumer, H (1969) *Symbolic Interactionism: Perspective and Method*, Englewood Cliffs Prentice-Hall.
- Freeman AJ, Chen YL (2019) Interpreting pediatric intelligence tests: A framework from evidence-based medicine. In: Goldstein G, Allen DN, DeLuca J (Eds.), *Handbook of psychological assessment*. Elsevier Academic Press, pp. 65-101.
- DeYoung, CG (2020) Intelligence and personality. In: RJ Sternberg (Ed.), *The Cambridge handbook of intelligence*. Cambridge University Press, pp. 1011-1047.
- Doerfler, T (2007) *Intelligenz, Persönlichkeit und die mediiierende Wirkung von Itembearbeitungszeiten*. (Doctoral Dissertation, Rheinisch-Westfälischen Technischen Hochschule Aachen.
- Eckert C, Schilling D, Stiensmeier-Pelster J (2006) Einfluss des Fähigkeitsselbstkonzepts auf die Intelligenz- und Konzentrationsleistung [The influence of academic self-concept on performance in intelligence and concentration tests]. *Zeitschrift für Pädagogische Psychologie* 20: 41-48.
- Steinmayr R, Wirthwein L, Schöne C (2014) Gender and numerical intelligence: Does motivation matter? *Learning and Individual Differences* 32: 140-147.
- Wechsler D (2003) *Wechsler Intelligence Scale for Children-Fourth Edition*. San Antonio, TX: Psychological Corporation.
- Dugbartey AT, Sanchez PN, Rosenbaum, JG Mahurin, RK Davis, et al. (1999) WAIS-III Matrix Reasoning test performance in a mixed clinical sample. *The Clinical Neuropsychologist* 13(4): 396-404.
- Roth, D (1978) Raven's matrices as cultural artifacts. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition* 1(1): 1-5.
- Sattler JM, Dumont, R (2008) WISC-IV subtests. In: JM Sattler (Ed.), *Assessment of Children: Cognitive Foundations*. La Mesa, CA: Sattler Publisher, pp. 316-363.
- Kubinger KD, Wolfbauer C (2010) On the risk of certain psychotechnological response options in multiple-choice tests: Does a particular personality handicap examinees? *European Journal of Psychological Assessment* 26(4): 302-308.
- Weiße RH (1998) *Grundintelligenztest Skala 2 (CFT 20)* [Culture Fair Intelligence Test, Scale 2]. (4th edn.). Göttingen: Hogrefe.
- Dowell LR, Mahone EM (2011) Perceptual Reasoning Index. In: Kreutzer JS, DeLuca J, Caplan B (Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY.
- Nisbett RE, Aronson J Blair, C Dickens, W Flynn, J Halpern, et al. (2012) Intelligence: New findings and theoretical developments. *American Psychologist* 67(2): 130-159.
- Steinmayr R, Spinath B (2019) Why time constraints increase the gender gap in measured numerical intelligence in academically high achieving samples. *European Journal of Psychological Assessment* 35(3): 392-402.
- Wechsler D (2014) *WISC-V: Technical and Interpretive Manual*. Bloomington, MN: Pearson.



This work is licensed under Creative Commons Attribution 4.0 License
DOI: [10.19080/GJIDD.2023.11.555825](https://doi.org/10.19080/GJIDD.2023.11.555825)

Your next submission with Juniper Publishers

will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats

(Pdf, E-pub, Full Text, Audio)

- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>