Response Processes as a Source of Validity Evidence in Self-Report Measures: Theoretical and Methodological Implications

by

Mihaela Sorana Launeanu

B.A., Psychology, University of Bucharest, 1998

M.A., Clinical Psychology, University of Bucharest, 1999

M.A., Counselling Psychology, The University of British Columbia, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE DOCTOR OF PHILOSOPHY in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES (Measurement, Evaluation, and Research Methodology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2016

© Mihaela Sorana Launeanu, 2016
Abstract

Response processes have been recognized as a primary source of validity evidence in testing by the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014) and others (Messick, 1989, 1995, 2000; Embretson, 1983, 2010). In spite of this role and scholarly support for this type of research (Bornstein, 2011; Borsboom, Mellenbergh & van Heerden, 2004; Zumbo, 2009), very few empirical studies have examined response processes as a source of validity evidence, and the theoretical and methodological grounding of this type of research has also been lacking (Padilla & Benez, 2014; Zumbo & Shear, 2011).

In response to these limitations in response processes research, this dissertation aimed to: (a) develop process focused theoretical models and coding systems for the response processes that underlie responding to self-report items, (b) apply the models to a well-known self-report measure of self-esteem (i.e., the Rosenberg Self-Esteem Scale; Rosenberg, 1965), (c) examine the theoretical and methodological implications of this investigation, and (d) propose several theoretical and methodological future directions in response processes research such as: expanding the scope of this research beyond individual cognitive processes towards socially situated response processes, exploring complementary theoretical and methodological foundations, and redefining the roles of response processes in validity research.

The findings of this dissertation will contribute to: (a) elaborating a process focused theoretical model of response processes associated with responding to self-report items, (b) adding to the empirical research data based on response processes, (c) enlarging the substantive validity basis of the self-esteem construct using a response processes method, and (d) proposing future theoretical and methodological directions in this area.
Preface

This dissertation represents an original, unpublished and independent work of Mihaela Sorana Launeanu. The research project presented in this dissertation has been approved by the UBC Research Ethics Board in October 2011, Ethics Certificate Number H11-02379.
# Table of Contents

Abstract ................................................................................................................................. ii  
Preface ................................................................................................................................... iii  
Table of Contents ................................................................................................................ iv  
List of Tables ....................................................................................................................... vi  
List of Figures ..................................................................................................................... vii  
Acknowledgments ................................................................................................................ viii 
Dedication ............................................................................................................................ x  
Chapter 1: Introduction ........................................................................................................ 1  
  Purpose of the Dissertation .............................................................................................. 1  
  Structure of the Dissertation ............................................................................................. 2  
  Response Processes as Substantive Validity Evidence ..................................................... 3  
  Response Processes and Contemporary Validity Theory ................................................ 8  
  Review of the Empirical Findings in Response Processes Research ................................ 14  
  Relevance of the Response Processes Research ............................................................. 24  
  Introduction ...................................................................................................................... 27  
  Conceptual Clarifications ................................................................................................. 27  
  The Building Blocks of a Response Process Model ........................................................ 31  
  A Model of Response Processes Underlying Answering Self-Report Self-Esteem Items .... 40  
Chapter 3: Examining the Substantive Validity of the Rosenberg Self-Esteem Scale Using Response Processes Inquiry ............................................................... 61  
  Introduction ...................................................................................................................... 61  
  Purposes and Research Questions .................................................................................... 61  
  Method ............................................................................................................................... 62  
  Data Analysis .................................................................................................................... 66  
  Results ............................................................................................................................... 68  
  Discussion ......................................................................................................................... 78  
  Contributions of the Study .............................................................................................. 99  
  Limitations of the Study ................................................................................................. 102
Chapter 4: Observations on Response Processes Research and Its Future Theoretical and Methodological Directions: An Essay ................................................................. 104

Introduction .................................................................................................................. 104
A Critical Evaluation of Response Processes Research ................................................ 104
Alternative Epistemological and Theoretical Grounding .............................................. 111
Methodological Horizons ............................................................................................. 118
Expanding the Scope of Response Processes Research ............................................... 127
Reimagining the Roles of Response Processes ............................................................... 136

Chapter 5: Concluding Chapter ..................................................................................... 144

Introduction .................................................................................................................. 144
Novel Contributions of the Dissertation Project .......................................................... 144
Limitations of the Presented Research .......................................................................... 148
Implications of the Presented Findings .......................................................................... 152

References ..................................................................................................................... 158

Appendix A: Poster for Participants’ Recruitment ......................................................... 181
Appendix B: Informed Consent ..................................................................................... 182
Appendix C: Personal Demographic Form .................................................................... 184
Appendix D: Sample of Possible Probes for Cognitive Interviewing ......................... 186
Appendix E: Mental Health Resource Sheet ................................................................. 188
Appendix F: Rosenberg Self-Esteem Scale (Rosenberg, 1965) .................................... 192
Appendix G: Theoretical Coding Scheme for Response Processes on the Rosenberg Self-Esteem Scale ........................................................................................................... 193
Appendix H: Response Processes Underlying Responding to RSES: Item Level Analysis .... 197
List of Tables

Table 2.1. Theoretical Underpinnings of the Self-Esteem Construct ..........................................................44

Table 2.2. Self-Processes and Response Strategies Involved in Answering the RSES Items........49

Table 2.3. Response Processes Underlying Responding to RSES Items.............................................53

Table 3.1. Response Processes not Anticipated in the Theoretical Model.................................71

Table 4.1. Reimagining the Roles of Response Processes.................................................................137
List of Figures

Figure 2.1. Observed Responses and Inferred Mechanisms: An Inductive View.........................29

Figure 2.2. The Building Blocks of a Response Processes Model........................................32

Figure 2.3. Integrative Response Processes Model.................................................................38

Figure 2.4. Situated Teleological Model....................................................................................42

Figure 2.5. Self-Processes Underlying Responding to Self-Report Items..............................47

Figure 2.6. Theoretical Response Processes Model for the RSES items...............................54

Figure 2.7. Situated Response Processes Model for the RSES items......................................57

Figure 3.1. Revised Response Processes Model for the RSES items.......................................98
Acknowledgments

First, I’d like to thank all the participants who took part in this research project, and to all clients whom I met during assessment or psychotherapy sessions. You inspired me to search for the meaning and the person beyond the test scores, and to seek to understand each of you beyond superficial categories or labels. I’ve deeply appreciated what you shared with me.

A special thank you to all my students, for being who you are, for all your questions, doubts, and searches. Each class you inspire me to become a better teacher.

A warm thank you, Maria and Alex for supporting me during my doctoral studies, and for putting up with the many hours that I spent away from you working on this project. Thank you to my parents for bringing me into this world, and for inspiring love and respect for education and knowledge. Multumesc, mama si tata pentru tot ce ati facut pentru mine!

I’d like to thank my dissertation supervisor, Dr. Anita Hubley, and to the members of my supervisory committee, Dr. Bruno Zumbo and Dr. Sheila Marshall, for the guidance and support that you offered during the sometimes challenging processes of writing my dissertation. I’ve very much appreciated your insights, feedback, and your trust in my ability to be successful in completing this project.

A heartfelt thank you to my colleagues from the MA Counselling Program at Trinity Western University: Dr. Marvin McDonald, Dr. Derrick Klaassen, Dr. Janelle Kwee, Dr. Bart Begalka, Dr. Rick Bradshaw, and Dr. Alex Kwee. Thank you for your unwavering support during this process, for your encouragements and sense of humor, and for standing by me. I am especially grateful to Dr. McDonald (‘Mac’) for the many hours spent dialoguing about alternative testing paradigms and contextualized assessment, and to my friends Derrick and
Janelle for the shared laughter, talks, walks, lunches and fun. Thank you Laurie, Shirley, Sarah and Robin for your amazing support during this journey, and for being such fine persons.

Last but not least, I am grateful for all the people- friends, colleagues or relatives- who believed in me, supported me, and were part of this journey in so many different ways. A big thank you to each of you!
Dedication

To my daughter, Maria, with all my love.

Te iubesc!
Chapter 1: Introduction

In this chapter, I will first introduce readers to the purpose and structure of this dissertation, and then I will provide a critical review of the current theoretical and methodological foundations of response processes research as well as of the empirical findings in this research domain. A description of the relevance of response processes research for validity and validation will conclude this chapter.

Purpose of the Dissertation

The overarching purpose of this research project is to contribute to the theoretical and methodological knowledge in the area of response processes as a source of validity evidence for the inferences based on test scores. Specifically, this dissertation project intends to be a proof of concept with respect to how to build and evaluate response processes models in testing, and proposes a way of theorizing about response processes as well as a strategy of model building in this area of research. To this end, this dissertation: (a) proposes a theoretically informed, model based approach to the study of response processes that underlie responding to self-report items, in general, and to self-report self-esteem items, in particular, (b) evaluates the proposed theoretical model using empirical data and makes suggestions for revisions of the model, and (c) examines future theoretical and methodological directions in the area of response processes research. In addition, this dissertation project makes an important contribution to the substantive validity evidence of a well-known and widely used self-report measure of self-esteem, the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). Although the RSES is the most widely used self-report measure of self-esteem, to my knowledge, no research studies have yet investigated the validity evidence for RSES inferences using response processes inquiry.
Structure of the Dissertation

This dissertation is structured in five chapters, as follows:

Chapter one introduces readers to the theme of responses processes in validity and validation by examining the theoretical and empirical state of art in response processes research. The theoretical underpinnings of response processes research are discussed in relation to contemporary developments in validity theory, and empirical findings in response processes research are reviewed.

Chapter two proposes a theoretical model of response processes underlying self-report self-esteem items. The model incorporates substantive knowledge from the areas of self, self-processes (i.e., self-evaluation, self-attribution, and self-regulation response processes), and self-esteem, and represents the theoretical framework for investigating the response processes underlying answering the items of Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), a well-known and widely used self-report measure of self-esteem.

Chapter three presents an empirical evaluation of the response processes model elaborated in the previous chapter and applied to Rosenberg Self-Esteem Scale items. The results obtained from interviewing research participants about their responses to the items of the Rosenberg Self-Esteem Scale are discussed both in light of the proposed theoretical response processes model and with respect to the contributions made to the validity evidence for the inferences based on the RSES scores. Revisions to the proposed model as well as the contributions and limitations of this study are addressed.

Chapter four provides a critical evaluation of the current theoretical and methodological foundations of response processes research, and, then, proposes and examines several new directions in response processes research such as: (a) exploring complementary theoretical and
methodological foundations, (b) expanding the scope of this research beyond individual cognitive processes and towards socially situated response processes, and (c) redefining the roles of response processes in validity research. Implications for validity in testing are discussed.

Chapter five, the concluding chapter of this dissertation, discusses the theoretical and methodological contributions of this research project: (a) elaborating a process focused theoretical model of response processes associated with responding to self-report items, (b) adding to the empirical research data based on response processes, (c) enlarging the substantive validity basis of the self-esteem construct using a response processes method, and (d) proposing future theoretical and methodological directions in this area. The limitations of this research project are also addressed.

**Response Processes as Substantive Validity Evidence**

Validity “refers to the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests” (*Standards for Educational and Psychological Testing*, AERA, APA & NCME, 2014, p. 11). Validity lies at the foundation of measurement and testing, as “…without validity, a test, measure, or observation and any inferences made from it are meaningless” (Hubley & Zumbo, 1996, p. 207). Response processes represent one of the primary sources of validity evidence for the inferences made from test scores, and they have been characterized as the hypothetical (theoretical) mechanisms underlying the observed test scores. Identifying the response processes associated with answering test items would help clarify or explain how the test scores were produced and, ultimately, what they mean. In this sense, Borsboom et al. (2004) have argued that no other strategy of collecting validity evidence would ever be able to substitute the validity evidence coming from understanding and explaining how test scores were generated. Zumbo (2009) has also argued that “validity is not established until
one has an explanatory model of the variation in item responses” (p. 71), and that a strong form of validity is “theory driven and provides an explanation of the test scores” (p. 69). Investigating the response processes underlying test scores seems to be at the center of this endeavour.

Explaining how test scores were generated implies the development of explanatory models of test score variation (i.e., response processes models) informed by the theory of the attribute targeted by the test (i.e., construct theory) as well as by the relevant epistemological frameworks and substantive knowledge. These explanations could, for instance, take the shape of a deterministic theory of response behaviour (Borsboom et al., 2004; Borsboom, 2005) or that of a pragmatic, contextual explanation (Zumbo, 2009). Ideally, these theoretical explanations of the mechanisms underlying test responses should be formulated before administering a test so that hypotheses about the expected response processes stemming from these theoretical explanations can be stated and then compared with the observed response processes during testing.

Messick’s legacy. The importance of studying the response processes that occur during testing was explicitly stated by Messick (1989, 1995, 2000), who considered response processes to be one of the six sources of construct validity and the primary source for supporting the substantive validity of the inferences based on test scores. According to Messick (1989), in order to establish a solid validity basis, "the most illuminating of all […] are direct probes and modeling of the processes underlying test responses […]. At the simplest level, this might involve querying respondents about their solution processes or asking them to think aloud while responding to exercises during field trials" (p. 743). Furthermore, Messick (1989) stated that "the substantive aspect of validity refers to theoretical rationales (i.e., why; explanations and mechanisms) for the observed consistencies in test responses, including process models of task
performance (Embretson, 1983), along with empirical evidence (i.e., from coding) that the theoretical processes are actually engaged by respondents in the assessment tasks” (p. 744).

If we pause and reflect on Messick’s (1989, 1995, 2000) thoughts on this topic, we notice that they point towards several important theoretical and methodological issues in the area of response processes research, and that they set the tone for what response processes are and how they should be investigated by researchers. The first and perhaps the most relevant remark is that Messick explicitly connected the empirical study of response processes with the investigation of the theoretical rationales underlying these processes, and with the development of theoretical models that may explain or account for the observed consistencies in response processes. This endeavour requires more than just interviewing the test takers about what they think when answering test items or when performing a task. The findings of these interviews, although very important, represent only one part of the substantive validity puzzle, namely, in Messick’s words, “the observed consistencies” in the test takers’ responses. The other significant part, that legitimizes the results of this type of investigation as substantive validity evidence, pertains to the elucidation of the theoretical rationales that explain or account for these observed consistencies. Providing substantive validity evidence means accomplishing these tasks in an integrated fashion.

A second observation that is important to note is that, although Messick (1989, 1995, 2000) proposed a method for studying response processes during testing tasks (i.e., asking test takers about these processes), he considered this method to be the simplest of all. Thus, other methods may offer additional or complementary insights to the information obtained through interviews or think aloud protocols (TAP), and these complementary methods await to be explored. And certainly he implied there are likely to be better methods through which one may
understand response processes. In this sense, Sireci (2009) mentioned other possible strategies of obtaining evidence about test takers’ response processes such as: “systematic observations of test response behaviour, evaluation of the criteria used by judges when scoring performance tasks and analysis of item response time data” (p. 18). Nonetheless, so far, the vast majority of studies in the area of response processes have used either cognitive interviewing or TAP to investigate these processes. Extremely little work has been invested in the exploration of alternative methods to study response processes or in the development of comprehensive methodological frameworks and guidelines that could direct and focus the study of response processes in testing.

Lastly, Messick's (1999) comments above were inspired primarily by his reflections on the research studies conducted in the domain of cognitive, performance-based testing, and, further in his comments, he quoted research studies conducted by Frederiksen, Mislevy, and Bejar (1993), and by Snow and Lohman (1989) to provide some examples of this type of inquiry. Messick’s reference to these studies is not surprising given that the most intensive work in the study of response processes has been accomplished in the domain of cognitive testing (Embretson, 2010; Embretson & Gorin, 2001). Notwithstanding, much less is yet known about the response processes that occur when answering self-report measures, for example. So far, the research on response processes has not yet differentiated between various types of tests and testing requirements, and this differentiation may become important for both theoretical and methodological reasons that will be addressed later in this dissertation.

The Standards for Educational and Psychological Testing. According to the Standards for Educational and Psychological Testing (AERA, APA & NCME, 1999, 2014), validity evidence based on response processes is one of five key sources for the validity of inferences based on test

---

1 The short reference Standards will be used to cite this document hereon in.
scores, and it refers to “evidence concerning the fit between the construct and the detailed nature of performance or response actually engaged in by examinees” (p. 12). Moreover, the Standards recommended that response processes investigation address “questions about difference in meaning or interpretation of the test scores across relevant groups of examinees” (p. 12). From these quotations, it seems quite clear that substantive validity evidence as seen by the Standards is mainly about investigating the fit between the responses expected based on the construct theory as outlined in the test specifications, and the response processes observed during testing and whether these are consistent across groups of examinees.

Although it is important to commend the Standards (AERA, APA & NCME, 1999, 2014) emphasis on the fit between theoretical rationales and observed responses, it is also timely to notice that the vast majority of widely used measures in psychology do not come with a manual or description specifying the expected response processes that test takers should demonstrate in order for test inferences to be valid. Therefore, practitioners engaged in validation studies of such measures cannot simply resort to test specifications in order to straightforwardly compare or match test takers’ observed responses or response processes on a test with the expected response processes elaborated in the test manual. Thus, although well-intended, some of these standards are difficult, if not downright impossible, to uphold in actual validation practice, and this reality has created some degree of confusion and frustration among researchers who have attempted to conduct validation studies using response processes.

Conclusion. Pondering Messick’s (1989, 1995, 2000) writings as well as the provisos outlined in the Standards (AERA, APA & NCME, 1999, 2014), several main points are important for the present investigation: (1) response processes represent the primary source of substantive validity evidence for inferences based on the test scores, (2) the mechanisms or
processes that underlie the observed consistencies or patterns in response processes that occur during testing have to be explained or modeled in light of the theoretical foundation of the construct of interest in order to provide sound substantive validity evidence; thus, just reporting the inventory of the observed response processes during testing is not enough for this purpose, (3) the primary focus of response processes research as substantive validity evidence is on elucidating the generative mechanisms (i.e., processes) underlying test takers’ responses, not just indexing the content of these responses, and (4) although interviewing test takers during testing seems to be the simplest method to access these response processes, there are still lots of methodological gaps with respect to how to best research these processes.

Building on these conclusions, I will next: (a) examine the theoretical underpinnings of response processes research in order to conceptually situate this type of research within the broader landscape of contemporary validity theory, (b) review and critically evaluate the main empirical findings in this area, (c) identify and characterize the main categories of response processes relevant to psychological testing and validation, and (d) examine the relevance of response processes research to testing, contemporary psychometrics, and validity.

**Response Processes and Contemporary Validity Theory**

Response processes play an important role in our contemporary view of validity. In response to Messick’s (1989, 2000) encouragements to devote more attention to the study of response processes as a primary source of evidence for substantive validity and, in line with provisos from the *Standards* (AERA, APA & NCME, 1999, 2014) that included the examination of response processes as one of the main sources of validity evidence, several scholars and researchers in the field of validity have further pursued the examination of response processes and their role in contemporary psychometrics and test validity. In this section, I will review how
response processes have been conceptualized in contemporary validity theories in order to delineate the current theoretical foundations of response processes research.

**Construct representation.** Embretson (1983, 2010) has commented explicitly and extensively about the essential role that response processes play in construct validity and in the context of model-based measurement, and about some strategies to study these processes. In this sense, she has made an important conceptual contribution to understanding response processes as an integral part of construct validity theory by distinguishing between the ‘nomothetic span’ of a construct (i.e., the relationships between test scores and an array of theoretically related variables) and the ‘construct representation’ or the theoretical mechanisms that underlie item responses/test scores (Embretson, 1983). For Embretson, elucidating construct representation means “to illuminate the processes in which people engage while completing psychological tests” (Bornstein, 2011, p. 534). Thus, in her conceptualization, Embretson has explicitly recognized the dynamic aspects of construct representation as an intrinsic part of the construct theory and, thus, of construct validity, as well as the critical role of response processes in designing tests that accurately target and represent the intended construct.

Moreover, Embretson (2010) believed that the “success of the construct modeling approach, especially for construct representation research, will depend on the ability of researchers and test developers to develop quantitative indices that define the theoretical mechanisms that are involved in the tasks” (p. 195). This objective highlights the role of response processes in psychometric modelling, model-based measurement, and test development: potentially, some of the identified underlying mechanisms of test takers’ responses could be quantified and estimated via psychometric models. Some of the underlying mechanisms that Embretson suggested and found especially relevant for the domain of cognitive testing are:
information processing, problem solving and decision making strategies, and the type of cognitive information or knowledge mobilized when answering test items.

To accomplish the above stated goals, Embretson has developed an active research program in the area of cognitive testing using an item response theory (IRT) psychometric approach. Hence, most of her practical insights apply best to this testing domain. In essence, Embretson (2001, 2010) made use of cognitive psychology principles, theoretical models, and empirical findings in order to improve test design and items in the domain of cognitive testing. For example, she linked item development with cognitive complexity as psychological process, and stated that “items can be developed to target difficulty levels with specified sources of cognitive complexity” (Embretson & Gorin, 2001, p. 265). Cognitive complexity demonstrated during various cognitive tasks was examined as part of the response processes research agenda, and was later modeled as an item characteristic and estimated using IRT methods. In so doing, Embretson and Gorin demonstrated that cognitive theory can be influential not only during the test interpretation phase but also during the item development phase. Thus, this type of work in the domain of cognitive testing represents a significant signpost about purposefully integrating theory (i.e., cognitive psychology theories) in building and refining test items, while using more advanced model-based measurement strategies such as IRT models. Notably, however, Markus and Borsboom (2013) commented that, although very valuable, these cognitive mechanisms are too abstract and removed from the actual meaning of test items, and thus they may not capture all of the nuances involved in responding to test items.

**Theories of response behaviour.** Borsboom et al. (2004), Borsboom (2005), and Markus and Borsboom (2013) considered response processes to be a central ingredient of a robust validity theory. Borsboom et al. advocated extensively for developing a "theory of response
behaviour" (p.75), which was a causal explanation for how the measured attribute produces test score variation. They also encouraged researchers to focus on identifying the response processes that would be a critical part of such a theory. In essence, the term “response behaviour” represents yet another name for response processes, and the theory of response behaviour resonates well with both Messick’s and Embretson’s views on developing theoretical explanations for these response processes.

Borsboom et al. (2004) regarded the theory of response behaviour as the only way to ensure test validity and, to this end, they stated that, if and only if one can formulate and test casual theories that explain how test score variations were generated by the construct that the test is measuring, can one then convincingly demonstrate that the test is indeed measuring what it is supposed to measure, and, thus, is a valid test. Whereas some of the epistemological assumptions of such an argument are disputable (e.g., the realistic ontology of the latent variables, and the strict, linear deterministic view of test score explanations), this endorsement of the development of theories of response behaviour represents yet another timely reminder of the importance of studying response processes in order to articulate coherent and integrated explanations about test scores and, ultimately, to strengthen the validity of the inferences made from these scores.

Pragmatic-contextual explanatory models. Zumbo (2007, 2009) has espoused a pragmatic and context based explanation (van Fraassen, 1980) as an alternative explanatory model for observed test score variation. Epistemologically, this type of explanation is more nuanced and open to contextual influences than Borsboom’s strictly deterministic explanatory model. Thus, this type of explanation may be more responsive to the dynamic and fluid reality of validity and validation. Determining the mechanisms of a pragmatic, contextual explanation of test scores allows one to uncover the response processes involved in generating, constructing, or
shaping the test scores not only at an individual level, but also at social and other trans-individual or cultural levels (e.g., school, neighbourhood, communities). Hence, the ability to formulate multilevel explanatory models represents an important virtue of pragmatic explanations.

In addition to these theoretical developments, Zumbo’s empirical research in the area of pragmatic, contextual explanation related to validity has focused predominantly on quantitative multilevel modeling (Forer & Zumbo, 2011; Ruhe & Zumbo, 2009; Zumbo & Gelin, 2005). In this sense, an important contribution was the development of the “third generation of DIF models” (Zumbo, et al., 2015) that allows one to include the context within statistical modeling via latent classes or mixture IRT models. Instead of assuming invariance across contexts, these newer models allow one to include contextual variables at the heart of psychometric modeling. This represents a significant stride forward in modeling and explaining response processes that generate test score variation, and a much needed alternative to the strictly deterministic explanations espoused by Borsboom et al. (2004).

**Experimental process based validity.** Bornstein (2011) is another researcher who has contributed to the discussion about the role of response processes in contemporary test validity. In Bornstein’s article about process based validity, he elaborated on some of Embretson's ideas, and suggested that substantive validity evidence stemming from response processes validity should be regarded as "process based validity" (p. 535). To this end, he defined process based validity as "the degree to which respondents can be shown to engage in a predictable set of psychological processes during testing. Once these processes are identified, experimental manipulations are introduced to alter these processes and determine whether the manipulations affect test scores in meaningful ways" (p. 536). Bornstein has also proposed a general framework for conducting process based validation studies. Although his suggestions are a practical step
forward in elaborating some of Embretson's ideas outside the cognitive testing domain, they are still very general and with yet unexplored applied potential. One of Bornstein’s merits is that he has brought explicitly to the fore the need to move away from strictly observational, correlational approaches towards more experimental, active approaches in identifying relevant response processes across a variety of testing domains and situations. This impetus resonates with what Markus and Borsboom (2013) identified as the necessary move from “look and see psychometrics” towards “try and see psychometrics”.

**Conclusion.** The study of response processes is well-anchored and well recognized within the theoretical landscape of contemporary validity. Response processes are seen as the primary source of substantive validity evidence (Embretson, 2010; Messick, 1995; Zumbo, 2009), as an intrinsic part of the construct, namely construct representation (Embretson, 1983), and as explanatory mechanisms of test score variation (Borsboom et al., 2004; Zumbo, 2009).

Some of the most important validity scholars have strongly advocated for developing response processes models as explanatory models of test scores (Borsboom, 2005; Zumbo, 2009) in order to strengthen the foundation of substantive validity for measures used in the common practice of testing. Epistemologically, the proposed theoretical models have ranged from strictly causal, deterministic explanatory models to pragmatic, contextual explanations. Embretson’s work (2010) exemplified the integration between foundational psychological theory and test development in the domain of cognitive testing where she has developed and implemented some local working models for explaining item responses on various cognitive tasks. However, a comprehensive theory or explanation of cognitive response processes involved in testing has not yet been generated (Markus & Borsboom, 2013), and discussions that have occurred about response processes models have been largely abstract and philosophical with very little
applicability in the concrete activity of testing and validation. This may be partly due to the inability of a ‘grand theory’ to account for local testing realities and models, and partly due to the significant theoretical and methodological effort required by such an endeavour. In sum, in spite of some theoretical controversies, response processes represent a fecund and promising area of research in contemporary validity.

**Review of the Empirical Findings in Response Processes Research**

In this section, I will review the main empirical findings in the area of response processes research, and I will: (a) discuss the prevalence of research studies examining response processes as a source of validity evidence, (b) describe the methods that have been used in these studies, (c) review the main categories of response processes that have been empirically investigated, (d) conduct an evaluation of these research findings and current trends in response processes research, and (e) discuss the limitations and potential of this type of empirical research.

**Prevalence of empirical research studies on response processes.** Overall, relatively few research studies have explicitly examined response processes as a source of substantive validity evidence. Zumbo and Shear (2011) conducted a review of representative academic journals and noted that, in the last 20 years, the number of publications about the search for evidence of response processes has hardly increased and that the traditional sources of validity evidence, such as convergent and discriminant validity or factor analytical methods, have clearly overwhelmed the timid attempts to provide substantive validity evidence via response processes inquiry. Moreover, Zumbo and Shear noticed that the highest presence of reported research findings on response processes has been found in the medical outcomes field, where 14% of validation studies were based on participants’ response processes. In other research domains,
very few studies have focused on collecting validity evidence based on test takers’ response processes.

The findings of Zumbo and Shear’s (2011) review are consistent with those of a previous review conducted by Cizek, Rosenberg, and Kong (2008). After investigating the sources of validity evidence reported in the *Mental Measurements Yearbook*, the authors concluded that response processes were only investigated for 1.8% of the measures, whereas criterion-related (correlational) validity evidence was provided for 67.2% of the measures. None of the personality/psychological measures or social measures reported on response processes as sources of validity evidence, whereas this type of evidence was reported in 5.9% of the cases of the developmental measures, 4.0% of the behavioural measures, and 3.7% of the achievement measures (p. 39).

The most recent review of response processes research studies (Padilla & Benitez, 2014) indicated that the large majority of empirical studies reporting response processes findings were validation studies in the area of health, and that this type of evidence was collected predominantly during the development of scales for specific diseases or health processes. For example, Deal, DiBenedetti, Williams, and Fehnel (2010) employed response processes inquiry to develop and validate a pain scale. Althof, Perelman, and Rosen (2011) used the same strategy with a sexual arousal scale, and Brod, Hammer, Christensen, Lessard, and Bushnell (2009) implemented response processes interviews for identifying patients’ perspectives about diabetes. Extremely few papers have focused on investigating psychological cognitive processes, such as processing speed (Cepeda, Blackwell, & Munakata, 2013) or decision making strategies (Day, 2010).
In sum, the findings of these reviews suggest that health and medical outcomes research fields have been the most likely to launch investigations of response processes as a source of substantive validity evidence. There has been extremely little to no interest in examining response processes involved in responding to psychological questionnaires and to investigating psychological processes engaged during various testing tasks. In research, a significant disparity between traditional, correlational validity, and response processes based validity has persisted and even flourished. This may be due to poor research habits, convenience, and, potentially, to the lack of realistic guidelines about how to conduct research on response processes in testing (e.g., the 2014 Standards’ view on response processes is largely inapplicable in the real world of validation given that most tests do not have a priori specifications about the expected response processes).

There is also a significant gap between the theoretical aplomb demonstrated by several scholars and researchers who have advocated for the importance of response processes research as substantive validity evidence and the actual reality of the empirical research on response processes. This gap suggests that validity theory does not easily translate into practice, possibly due to its abstract nature and unclear practical guidelines. Thus, it is unlikely that change in the empirical status of response processes research will come from the debates among the theoreticians of validity, but rather from the field work of researchers who will find a way to integrate the suggestions outlined in these theories with the concrete, applied realities of their practice. So far, there are few examples of such an integration in the empirical research on response processes (e.g., Embretson’s work) whereas in other research domains, particularly the social-personality field of psychology, there has been little to no interest regarding the examination of response processes.
Methods used in research studies on response processes. Compounding the obvious scarcity of research studies in this area, a concerning aspect noticed in all of the reviewed studies pertained to the methods that the researchers have used in order to collect validity evidence via response processes. The overwhelming majority of the surveyed studies used cognitive interviewing or a modified version of it in order to collect evidence about response processes (DeWalt et al., 2007; Ercikan, Arim & Law, 2010; Gadermann, Guhn & Zumbo, 2011; Irwin, 2009; Olt et al., 2010; Poole, 2009; Willis, 2005). Extremely few studies have used other methods to investigate response processes, such as eye tracking and response times as indirect indicators about response processes.

Moreover, in a candid spirit of both methodological blindness and an atheoretical outlook, the vast majority of the reviewed studies asked respondents about what they thought while completing a test without having any a priori hypotheses about what this inquiry might reveal. The end result of this enterprise was the notable absence of a convincing explanation for the response processes investigated, together with a minimal and superficial connection between the empirical findings and the theory of the construct targeted by various tests.

Overall, this suggests that there is a concerning scarcity of methods and methodologies employed in the area of response processes research. This very limited methodological outlook may be partly due to the absence of relevant guidelines about how to conduct response processes research as well as because of not exploring alternative methods and methodologies in the empirical investigations of response processes research.

Categories of Response Processes. The review of the empirical research studies on response processes suggests that, by far, the most researched response processes were cognitive processes, such as item comprehension, interpretation, retrieval, and expression of an answer to
test items (Schwarz, 1999; Tourangeau, 1994). In addition, Embretson (1983, 2010) also identified information processing and problem solving strategies among the relevant cognitive response processes, and Embretson and Gorin (2001) identified cognitive complexity and the cognitive operations involved in dealing with cognitive complexity as central to modeling item difficulty level using IRT. These processes are highly relevant to cognitive testing and represent important milestones towards developing a response processes model in the cognitive assessment domain, but they do not necessarily represent the most relevant response processes on other testing domains or with respect to other types of tests (e.g., self-report measures).

Given the narrow focus on cognitive processes, most of the existing guidelines about how to study response processes have recommended cognitive interviewing as the method of choice to tap into these processes (Wills, 2005). In spite of this sustained focus, a comprehensive theory of cognitive response processes that underlie test ratings has not yet been proposed and tested (Markus & Borsboom, 2013). It is disputable whether such a grand theory is even possible or desirable given that the “one size fits all” approach is unlikely to inspire future research studies on response processes.

Although there is no doubt that cognitive processes such as understanding and interpreting test items are critical and are likely to be involved to a certain degree in all testing activities (e.g., any test item poses a cognitive load), cognitive processes are not the only processes involved in answering test items. The over-focus on researching cognitive processes may stem from a combination of methodological problems (e.g., cognitive interviewing as the most used method in response processes research), historical-contextual considerations (e.g., “cognitive friendly” research environments), and some outstanding achievements in the cognitive testing arena that are not yet matched by similar accomplishments in the non-cognitive...
domains. Future empirical research should broaden its focus to include empirical investigations of non-cognitive response processes such as: self-referential, motivational, affective, situated, and dialogical response processes, to name but a few. It would be desirable if programs of research could be elaborated upon, with the purpose of addressing these non-cognitive response processes as a significant source of substantive validity in testing as a complement to the focus on cognitive processes.

**Evaluation of the empirical findings.** In spite of the new information that research studies on response processes have brought to validity table discussions, in most cases, research data stemming from response processes studies have not been yet systematized or connected with the theoretical rationales underlying the construct targeted by different measures. Even the reviews conducted in this area have limited their scope to mere descriptive statistics (e.g., counting how many articles addressed response processes, and in what field), rather than critically evaluating the importance of the findings. Therefore, it is hard to evaluate whether, and to what extent, these findings have impacted validation practices or the development of better measures. It is likely that, overall, the impact has been minimal, with the exception of some local research programs. With few exceptions (Embretson, 2010), the empirical research on response processes seems to be largely empiricist, purposeless, and disconnected from its theoretical moorings. Nonetheless, in the midst of this, there are some distinct tendencies worth noting and evaluating.

**Methodological myopia.** If we conceive of our methods as our lenses into the study of a phenomenon, the conclusion of these empirical reviews is quite concerning: we have really only looked at response processes through one single pair of lenses while trying to capture different aspects of a broader and multilayered validity landscape. Thus, we might have gotten a very
blurry or distorted picture in some ways, and we might have also missed many relevant angles and nuances.

The over-representation of cognitive interviewing or TAP as basically the “gold standard” methods to study response processes may be due partly to the lack of theoretical rationales behind conducting such studies, and partly it may reflect a rather simplistic understanding of response processes. As mentioned previously, exploring response processes as validity evidence is not only about asking test takers what they think while they answer test items. Rather, it requires a purposeful development of integrated theoretical and methodological frameworks, and then a rigorous application of these frameworks to researching the observed response processes. Given this severe methodological myopia, it becomes essential for this type of research to use some “prescriptions” in the form of comprehensive methodological frameworks and realistic guidelines that address how to study response processes as substantive validity evidence. For example, Mislevy, Steinberg, and Almond (2002) developed a model-based reasoning for the development of tasks and items, which may offer a practical example of researching and including response processes in the development of measures. This kind of applied work is very important given that the most notable absence is on the decision of how exactly to study response processes and how to incorporate them in validity arguments (Padilla & Benez, 2014).

Lack of theoretical grounding. Another very concerning tendency apparent from reviewing the empirical studies on response processes was their primary or exclusive focus on surveying respondents’ thoughts, impressions, or decision making strategies observed during testing without elucidating the theoretical aspects of these response processes or the “theoretical rationales” (Messick, 1995) of these observed response consistencies. Sadly, this lack of
theoretical foundation has been a longstanding, chronic problem in psychometrics and testing and, unfortunately, the research in the area of response processes makes no exception. However, it is unlikely that a simple index of TAP contents will produce any substantial validity evidence or that it will significantly contribute to understanding response processes, let alone to developing response processes models. In order to provide sound substantive validity evidence, theoretical models from psychology and other social sciences have to be mobilized in conjunction with the empirical findings collected via cognitive interviews, TAP, and, hopefully soon, other methods.

In addition to the scarcity of theoretical investment, the notable absence of response processes research in psychology and the social sciences, and the important absence of empirical research focused on investigating psychological response processes in other fields is also perplexing. The latter, in itself, may partly explain the methodological scarcity and the perpetual confusion between content and process in response processes research. These observations bring to the fore the comment about the shocking disinterest of psychologists in the concepts of their own discipline (Markus & Borsboom, 2013) as well as Kazdin’s (2007) suggestion addressed mainly to clinical psychologists to stop accumulating piecemeal correlational research and to start a purposeful exploration of the mechanisms, processes, and specific pathways of psychological change.

**Over-focus on response contents and strategies.** Another prevalent tendency in response processes research is the over-focus on *content* at the expense of examining the processes that underlie test responses. If we want to improve substantive validity and not simply transition from quantitative inventories of correlation matrices packed as validity evidence towards semi-quantitative indexes of frequency tables of how many times a participant has used certain words
or strategies during the TAP sessions, then it is important to emphasize the study of response

processes and not just response contents. It is doubtful that only counting test takers' response contents alone will produce any significant contributions to the substantive aspect of test validity. At its best, response content analysis may provide some additional measure of content validity but, in the long run, it may become cumbersome, difficult to manage, and eventually it will drive researchers away from attempting to study response processes. What is needed in the field of response processes research is an epistemological development of process based models of test responses rather than indexing the contents and/or test takers’ strategies to arrive at these contents.

**Exclusive focus on cognitive intra-individual processes.** Research on response processes has been almost exclusively focused on examining the cognitive processes that occur during testing. In the cognitive testing domain, some processes have been identified and systematically researched. Some guidelines about how to investigate these cognitive processes have been developed and published (Wills, 2005). Nonetheless, a multitude of equally relevant response processes that shape test scores and their interpretations have been ignored by researchers. This persistent tendency reveals a very narrow landscape in the current research on response processes, and the existent research on response processes seems to have barely scratched the potential of response processes for test validity.

**“One-size fits all” approach.** Lastly, there has been a smooth and quite nonchalant shift from exploring response processes in cognitive, performance based testing to investigating response processes during self-report testing, without fully acknowledging the fundamentally different nature of self-report measures in terms of the response processes recruited by this method of assessment. For example, self-report measures involve a different kind of
introspection than that required to describe to someone the steps taken in solving a problem, and, thus, they may involve qualitatively different response processes such as: reflexivity, self-construal and self-evaluation strategies, meaning making processes, etc. While using modified forms of TAP, sometimes in conjunction with methods such as eye tracking and reaction times, has generated significant strides forward in the area of cognitive testing and performance assessment (Embretson, 2010), the progress is unclear or has not yet been convincingly documented when using TAP to investigate self-report response processes. This may be partly due to the scarcity of response processes research in the field of psychology and social sciences in general, as well as to the lack of research of psychological processes other than cognitive ones. This reality requires a careful examination of the nature and specificity of response processes in self-report measures in order to determine the most appropriate method to study them, and, even more importantly, to decide how to use this rich knowledge to create better measures - which is, arguably, the purpose of studying response processes.

**Conclusion.** The review of the empirical findings in the area of response processes has indicated that, although very good work has been completed in the cognitive testing domain, several limitations are prevalent and worth noting: (1) most research studies have focused on reporting inventories of observed response processes without thoroughly connecting them with theoretical rationales, (2) most published tests do not have specifications regarding theoretically expected response processes, and, thus, the observed response processes cannot be compared against such specifications, (3) most research in the area of response processes has been accomplished in cognitive testing whereas other test types (e.g., self-report) have received significantly less attention, and (4) the methodological scarcity of reviewed empirical studies has
limited the access and study of response processes beyond cognitive processes and problem solving strategies.

**Relevance of the Response Processes Research**

In spite of these limitations, response processes research is needed more than ever in order to improve the substantive validity evidence of existing tests and to inform the development of better measures. There is a lot of yet unexplored potential in this type of research that makes it very promising for future investigations and theoretical elaborations. Notwithstanding, empirical research in this area is still just beginning and researchers are struggling with a multitude of yet unclarified aspects.

A *de facto* lack of focus on investigating these processes as well as the absence of well-developed methodological frameworks in this area, has prompted some authors to openly express their reticence towards incorporating response processes during scale development activities or as part of the validity arguments. For example, Norman (2003) suggested that trying to understand the motivation and mental processes by which a person came up with a numerical answer on a self-report measure is, if not downright irrelevant for some research purposes, at least cumbersome from the point of view of a quantitative science because taking into consideration these individual mental processes “will ultimately increase error of measurement and add complexity to interpretation” (p. 246). Moreover, he stated that “using scales based on individualized questions in studies where results will be averaged across patients may, in the end, be counterproductive” (p. 247).

Whereas Norman’s (2003) criticism draws attention to the need to find ways to streamline the methodology for researching response processes, his comments occlude some very important issues in assessment and validity, and reflect a narrow view on testing and
validity. For example, although he may be right that (a) studies focused solely on group level differences may be explicitly not interested in exploring individual response processes, and (b) mathematically, the numbers obtained at the end of a measurement process may indeed average across individuals, the personal characteristics or experiences that we hope those numbers represent are much more difficult to average across individuals. This difficulty is compounded by the fact that, sometimes, we do not know if indeed what we think these numbers represent in order to average them is, in fact, so. Maybe here is a good place to remind ourselves that, as Zumbo (2009) stated, “It is rare that anyone measures for the sheer delight one experiences from the act itself. Instead, all measurement is, in essence, something you do so you can use the outcomes” (p. 66), and something that has social impact and consequences.

In the area of social and health sciences, numbers obtained from self-report ratings are commonly understood to represent human characteristics of some sort and are actively used to make decisions that ultimately affect peoples’ lives and their access to treatment, personal freedom, opportunities, or various forms of social support. From this point of view, it becomes not only a scientific requirement but also an ethical obligation to ensure that results obtained using self-report measures represent what they were intended to represent. Exploring and understanding response processes involved in answering self-report questions represent a necessary step towards the enhancement of both the substantive evidence of, and consequences related to, the validity of self-report measures (Hubley & Zumbo, 2011).

One of the first steps that would allow one to begin to address some of these difficulties is to bring in some methodological order by developing a conceptual framework for investigating response processes. The next chapter will contribute to this goal by proposing and discussing a generic framework for a response processes model that could be further elaborated and tested via
empirical studies. This generic model will be later concretized in the area of self-referential self-evaluative response processes with a direct application to the Rosenberg Self-Esteem Scale.

Introduction

Most of the difficulties and setbacks encountered in the empirical research on response processes stem from a combination of factors, such as: (a) unclear or insufficient understanding of what constitutes response processes in testing, and a long-standing confusion between the observed and theoretical aspects of response processes, (b) the pervasive absence of adequate theoretical frameworks or templates for building response process models that could be further empirically evaluated or tested, (c) insufficient familiarity with the types of response processes that should be expected during testing, as directed by theory of the construct, test type and other relevant factors, and (d) methodological limitations (e.g., not knowing how to research or access these response processes and automatically resorting to cognitive interviewing as well as lack of realistic, applicable guidelines about how to conduct this type of studies).

In an attempt to address some of these concerns, in this chapter I will: (a) discuss some important conceptual clarifications with respect to a correct understanding of response processes, (b) outline the necessary building blocks involved in developing a response processes model, and (c) elaborate and concretize this model template by applying it to responding to self-report self-esteem items (i.e., the Rosenberg Self-Esteem Scale, RSES; Rosenberg, 1965).

Conceptual Clarifications

In the next paragraphs, I attempt to bring some clarity with respect to what constitutes response processes as well as to conceptually differentiate among observed responses, theoretical response processes, and a response processes model, respectively.
**Observed versus inferred.** Response processes refer to “the theoretical mechanisms that underlie item responses” (Embretson, 1983, p.179), and to the processes in which test takers are expected to engage while answering test items (Bornstein, 2011). Borsboom et al. (2004) and Zumbo (2009) suggested that response processes represent the explanatory mechanisms of test score variation, causal or contextual, respectively. It is important to note that, in these definitions, a response process refers to something that is inferred from the observed patterns of test responses, and explained by relevant theory, and which represents a dynamic, unfolding mechanism (i.e., has a temporal structure concretized in discrete stages or pathways).

This observation is critical given that one of the most prevalent confusions in the empirical research on response processes has been that the same term, ‘response processes’, has been used interchangeably to denote two fundamentally different layers of a phenomenon: 1) the observed responses (e.g., test scores or interview comments) or the observed patterns of variation in test scores, and 2) the underlying or inferred response processes which represent the mechanisms underlying the observed test responses. Most empirical investigations that have asked the participants what they were thinking about while answering test items, reported these results as response processes evidence. Therefore, most of the empirical studies on response processes have been confined to reporting the patterns or regularities in the observed responses as response processes evidence instead of investigating the mechanisms underlying these observed responses. This propensity to conflate these two different levels has been further compounded by the documented test takers’ difficulty to introspect and directly report on their mental processes during think aloud protocol tasks. Almost always, when asked what was going on for them when answering test items, respondents automatically provide a content-based answer, not a process focused answer. Therefore, the latter has to be inferred and theoretically
explained via theoretical models and substantive knowledge. Figure 2.1 aims to illustrate the difference between observed responses or response patterns (i.e., the data collected during testing or assessment), and the underlying response processes or mechanisms.

![Diagram: Theory, Test Scores, Observed Response Patterns, Inferred Response Process]

**Figure 2.1. Observed Responses and Inferred Mechanisms: An Inductive View**

**What, how and why.** Observed testing or assessment data (e.g., either quantitative data in the form of test scores or qualitative data in the shape of interviews or respondents’ comments about the quantitative test scores) tend to organize in specific patterns or configurations that may suggest the presence of underlying generative mechanisms that could explain the variation in the observed test scores. Depending on the level of analysis of the observed data (e.g., individual, interpersonal, or contextual), one may identify coherent observed patterns in the data either at the individual level or at level of the interactive or emergent observed responses that may point towards the existence of meaning making or contextual processes that underlie these observed responses. In sum, the observed responses contribute to elucidating the content or the “what” of test takers’ responses to test items or interviews.
A step further, response processes represent the theoretically explained mechanisms or processes that underlie the observed responses on a test. If we worked inductively and post-hoc (e.g., after the test is built and used), and we start with the observed test scores collected during an assessment process, then we could infer the presence of the response processes and explain how they work by mobilizing relevant theoretical assumptions and models. Figure 2.1 is illustrative for this inductive, post-hoc analysis where test scores are organized into observed response patterns from which underlying mechanisms may be inferred.

If we specified the mechanisms that we expected to underlie answering test items before the test items are built (a priori) then we would consider the observed response patterns as concretizations of these underlying mechanisms, and then the arrows in Figure 2.1 would be directed from the underlying mechanisms to the observed responses. Ideally, for strong validity evidence, the model of the theoretically expected response processes would be specified a priori (before the test is administered). In practice, extremely few tests even mention response processes as source of validity evidence and, to my knowledge, there are no a priori models of response processes.

These underlying mechanisms may be construct specific (i.e., uniquely or at least predominantly associated with a certain construct), and construct non-specific or more general socio-psychological processes recruited by the construct (e.g., self-processes involved in responding to self-report items). For example, although self-assessment and self-evaluation as self-processes may be mobilized during responding to most self-report measures, they may play a specific, central role in explaining the scores on a self-esteem measure such as the Rosenberg Self-Esteem Scale. Similarly, memory represents a general, fundamental psychological process involved in answering self-report items sampling diverse psychological constructs (e.g., self-
esteem, depression, extraversion or traumatic events). However, how exactly memory processes are involved and what kind of memory is mobilized while answering specific test items (e.g., autobiographical versus declarative) may be different across different tests, and this differential involvement of different memory processes may be relevant to the interpretations made of test scores. Using a metaphor, we can visualize the more general, construct non-specific processes as the bedrocks through which many streams of water (i.e., constructs) may run. Identifying the relevant response processes underlying answering test items help answer the question “how” test takers come up with their answers on a test.

One of the main purposes of identifying response patterns and processes is to be able to integrate this knowledge in a coherent, testable explanatory model of test score variation. In this sense, a response processes model represents a theoretical model embedded within a larger epistemological view that explains or describes response processes underlying test score variation by using existing theoretical frameworks and/or substantive models and knowledge from various disciplines. Usually, a model requires theoretical intra- and interdisciplinary integration. A response processes model answers the question “why” test takers’ endorsed certain items or response options during testing.

**The Building Blocks of a Response Process Model**

Developing a response process model takes place over several steps and involves the elaboration of the following building blocks: (a) the epistemological foundation of the model, (b) the theoretical underpinnings (i.e., the theory of the construct targeted by the test and relevant substantive knowledge), (c) the situated or contextualized dimensions of the model, and (d) the integrated model (i.e., the theoretical and contextualized dimensions integrated with the
epistemological tenets). These building blocks are presented in Figure 2.2, and each of these model constituents will be discussed in the next paragraphs.

\[Figure 2.2. \textit{The Building Blocks of a Response Processes Model}\]

\textbf{Epistemological foundations.} Any theory or theoretical model rests on an explicit or implicit epistemological foundation that circumscribes the assumptions, inferences and interpretations that may stem from applying the model to a particular field of knowledge. With respect to elaborating a response process model, clarifying the epistemological tenets of such a model represents a distinct meta-cognitive effort to provide a coherent and epistemologically well-anchored account of the test score interpretation. For instance, several epistemological stances have been proposed so far with respect to response processes in testing, such as: deterministic explanatory models (Borsboom et al., 2004), contextual-pragmatic explanatory models (Zumbo, 2009), and process-experimental models (Bornstein, 2011).

\textbf{Deterministic explanatory models.} Borsboom et al. (2004) proposed a deterministic model of response processes consisting of three components or levels: (1) the test scores as the
top/surface level, (2) the response processes as the intermediate level, and (3) the attribute measured by the test as the bottom level (i.e., the latent variable targeted by the test). In this model, response processes mediate between the attribute and test scores, and represent specific rules that translate the hypothesized (theoretical) attribute variation in the observed test scores variation. Response processes seen as rules aim to differentially connect the hypothesized variations in the attribute structure with the observed variations in the test scores. Thus, response processes represent differentiated, micro-level processes that are differentially activated by different levels or categories of the measured characteristics. In other words, response processes are flexible, dynamic rules that are turned on or off depending on the specific aspects of the measured attribute. Under this view, the translational rules are causal-deterministic, and the test is considered valid if and only if there is evidence that the attribute has causally produced the test scores.

Although this causal-deterministic model is enticing through its simplicity and clarity, it seems to be quite removed from the reality of testing, and thus, it could be, at best, a template for a formal psychometric model but not a model of response processes that occur during ‘real life’ testing situations. Moreover, the strict casual epistemological assumptions as well as the realist ontology that underlie this model are disputable on many levels.

**Pragmatic-contextual explanatory models.** Zumbo (2009) proposed a contextual-pragmatic explanatory model of response processes in which the context represents a distinct and important level of the explanation of test scores variation. Therefore, contextual variables (e.g., psychological, sociological, linguistic) are constitutive ingredients of the model. By introducing the notion of context, the arid and abstract picture proposed by the deterministic models turns into a more alive, dynamic and fluid picture, attuned to the larger context of testing.
Validity defined as the “inference to the best explanation” (Zumbo, 2009) in line with the pragmatic-empiricist view espoused by van Fraassen (1980) means that a unique deterministic relationship from attributes to test scores is not expected, and that, depending on the context, different explanations at different levels of generality can be employed. In other words, there is not only one way to explain the test scores but only a “best” way considering all the relevant variables, including the contextual ones.

**Process based experimental models.** A distinctive feature of the process-based experimental model suggested by Bornstein (2011) is its focus not only on identifying the relevant response processes underlying the test scores but also on experimentally manipulating various characteristics of the test items and potentially of the testing situation in order to identify more nuanced and context-dependent response processes. This type of model is valuable because it would allow one to experimentally explore how different response processes are differentially involved in the generation of the test scores. This is in line with the approach considering response processes as rules that are turned on/off in specific situations. Understood this way, the process-experimental model would also fit well with contextual explanations (Zumbo 2009), given that contextual variable could be an easy target of various experimental manipulations. Although Bornstein did not explicitly identify the epistemological assumptions of his model, his discussion about the experimental method and his references to Embretson’s approach suggest that he may be more inclined to support a causal process model grounded in experimental designs. However, an enhanced process-experimental approach could, in fact, illuminate significant situational or contextual aspects of response processes, and thus, it may fit well with contextual–pragmatic epistemological stances.
**Conclusion.** In spite of these different epistemological models being discussed in the contemporary validity debates, no fully elaborated and concretized response processes models have yet been proposed or tested. The discussions in this domain have been largely confined to a theoretical level with little impact on the practice of validation. This limitation may be partly due to the absence of methodological guidelines about how to build such a model and the following sections of this chapter will attempt to address this difficulty by expanding the inquiry from examining the epistemological foundations of such models to discussing other necessary ingredients such as the theoretical underpinnings, and the situated dimensions of the model.

**Theoretical model.** Another critical building block of a response processes model is to elaborate the theoretical, process-based underpinnings of the model. The purpose of this step is to establish the theoretically expected processes or mechanisms that underlie the construct of interest, both the construct specific processes and the relevant more general psychosocial processes recruited by the construct (e.g., self-processes, memory processes). Both the general and construct specific processes actively shape the test scores, and, ultimately the interpretations based on these test scores. For example, in the case of the self-esteem construct investigated in this dissertation, the following research domains and theories will be surveyed and integrated as part of developing the model: (a) specific mechanisms pertaining to the construct of self-esteem, and (b) self-processes, self-related emotionality and cognitions as general psychological processes. The process of integration between (a) and (b) allows to define the construct in terms of the underlying specific and general mechanisms, and, depending on the adopted epistemological stances, this model may explain the observed test score variation in terms of a differential involvement of these underlying mechanisms or response processes.
**Situated or contextualized model.** If we agree that testing is essentially a social and interpersonal activity and situation (Ardilla, 2008; Zumbo et al., 2015), then a comprehensive response processes model has to take into account the situated dimensions that shape the test takers’ responses to test items. In other words, test takers’ responses are not only a reflection of the theoretically expected responses and are not exclusively determined by the construct theory (e.g., strict causal explanation), but these responses represent an interactive blending between the theoretically expected processes on the one hand, and the contextual influences on the other hand. Instead of seeing context as a threat to validity as per the *Standards* (AERA, APA, & NCME, 2014), a situated model of response processes intentionally and purposefully integrates the context in explaining the test score variation (Zumbo et al.). Thus, the situated response processes model represents a contextualized, interactionist, and dialogical model that aims to explain test score variation in the context of a specific testing situation and task, and situated in a specific socio-cultural time and environment.

A useful epistemological framework for this type of situated model could be the ecological approach elaborated by Bronfenbrenner (1979, 1994), Fox (2001, 2003, 2004), McNamara (1997, 2007), and discussed by Zumbo et al. (2015), or the contextual empiricism described by Longino (1990). All of these scholars have emphasized the shift from focusing on test takers as individuals to focusing on the test takers within a testing situation, and how the interaction between test takers and the testing situation shapes the responses to test items. These situated and interactive response processes constitute a new level of response processes analysis: the four dimensional person-process-context-time proximal system (Zumbo et al.), and these response processes are an intrinsic part of a response processes model. In addition to the interactionist, *interpersonal* response processes, the contextualized model of response processes
should include socio-cultural factors such as how the social meanings and interactions, linguistic aspects as well as the larger cultural context shape the test scores and the interpretations made of these.

**Integrative response process model.** The integrative model would articulate the connections between theory of the construct and the context of testing within the proposed epistemological framework. Synthesizing nomothetic and idiographic frameworks would also pertain to this process of integration. *Figure 2.3* illustrates the embedded multiple layers of a response processes model, with the theoretical and contextualized or situated components integrated within a coherent, unifying model supported by the appropriate epistemological tenets that allow to provide plausible explanations and rich interpretations of test responses.
**Conclusion.** A comprehensive model of response processes that occur during testing represents a highly demanding and not always comfortable integrative effort across research domains and paradigms that may be naturally in tension. For example, balancing the hard-core epistemological assumptions of the psychometric models with the interpretative, hermeneutic focus of the situated validity discourse poses important conceptual and practical challenges.

Synthesizing the rational deductive, top-down approach of theoretically expected processes, the inductive, situated, and “bottom-up” approach of the contextualized response processes, and the interpretative dimensions of test scores requires generous integrative frameworks that are able to hold the tension that is naturally occurring between so-called “incommensurable paradigms”.

*Figure 2.3. Integrative Response Processes Model*
A response process model represents a multilayered explanatory-interpretative framework that integrates information from several sources, and should be able to allow the formulation of theoretical hypotheses about expected response processes concretized methodologically in theoretically driven coding schemes to be used during the empirical investigations in this area. Conducting extensive interviews about what respondents think during testing without having a hypothesis and with no clear direction, will only lead to large amounts of data that are cumbersome to understand and systematize. Ideally, every research study that aims to study response processes should outline a set of hypotheses regarding the expected mechanisms and response processes based on the theory of the construct, and on the relevant substantive knowledge in the domain of the investigated test.

A response processes model should be differentiated enough and tailored to the type of test, test purpose, and context. This means developing local hypotheses for different test formats (e.g., self-reports versus performance testing), resisting the automatic defaulting to cognitive interviewing in investigating response processes, and exploring alternative methods that may fit better with the construct, and with the test under study. Also, a sound model should account for response processes beyond the cognitive domain, as relevant (e.g., motivational, social, situational, dialogical).

A response processes model accesses and mobilizes relevant substantive knowledge and theoretical frameworks to develop hypotheses about the expected response processes, and is flexible enough to allow for ongoing experimentation, refinement, and revisions. Engaging in a process of mutual dialogue between theoretical rationales and empirical findings is necessary in order to come up with more refined construct theories and better items. Finally, a response
processes model should link and integrate response processes research findings with the rest of substantive knowledge within a research domain, and across multiple research areas.

**A Model of Response Processes Underlying Answering Self-Report Self-Esteem Items**

This section will elaborate the four building blocks of a response process model presented in *Figure 2.2* to propose a response process model underlying responding to the self-report self-esteem items of the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). Following the structure of the response process model illustrated in *Figure 2.2* and *Figure 2.3*, this section will discuss: a) the epistemological foundations of the model, (b) the theoretical underpinnings of the model, (c) the situated-contextual dimensions of the model, and (d) the integrative model of response processes underlying answering the RSES items.

**Epistemological foundations.** The model that I propose in this dissertation represents an epistemological blend of epistemic, teleological, and socio-cultural contextualism. The model is focused on the person in context, and represents an explanatory-interpretative model meant to provide coherent explanatory hypotheses regarding the observed scores variation on the RSES as well as a framework for the socio-culturally situated interpretations of the RSES scores.

**Epistemic and socio-cultural contextualism.** These epistemological lens will allow me to explain and interpret the RSES test scores as socially and contextually constructed in a mutual dialogue with the socio-cultural variables (e.g., language and socio-cultural meanings), and not as mere causal by-products or reflections of an abstract construct. As noted by Suarez (2014), “whereas the truth of causal claims is not necessarily relative to any context, the evidence for or against that causal claim is heavily dependent on the context of inquiry. The same observations, under the same interpretation, resulting from the same operations, performed in identical manner, will constitute evidence for a causal claim in one context but not in another” (p. 405). Thus, the
proposed model is contextual in the sense of epistemic contextualism that posits that the context determines the context of justification of knowledge. This epistemological stance is congruent with the social and interpersonal nature of the self-esteem construct (Rosenberg 1965; Leary & Down, 1995). In other words, I expect that the RSES scores are shaped by socio-cultural factors, and, therefore, the interpretations of these scores should these factors into account.

**Teleological model.** Test takers are already embedded in a context when they encounter the test items and the testing situation. They have not only a past but also desires, goals, and an imagined or perhaps preferred future, and, to a certain extent, expectations, hopes, and fears related to testing and to a specific testing situation. In this sense, the person comes before the test items, and he or she shapes how test items are perceived, understood, interpreted, distorted or dismissed. This dynamic interaction between the person and the items leads to decisions about how to respond the test items and these decisions are concretized in the test scores. Therefore, the model that I propose is teleological and anchored in test takers’ lived intentionality. Rather than a Stimulus-Person-Response model, it is a Person-Stimulus-Person-Response model (Figure 2.4). With respect to the validity of RSES scores, this teleological perspective means that an accurate and meaningful interpretations of test scores would take into account test takers’ intentionality (e.g., hopes, wishes).
Person-focused model. The proposed model rests on the basic observation that test scores are generated by the test takers, and not by abstract constructs or attributes. Therefore, in order to understand or explain how test scores are produced and what they mean, the focus of the inquiry will be on the test takers who answer the test items in a specific testing situation, and for a specific testing purpose. This epistemological position is consistent with the methodology used in this dissertation: conducting in-depth interviews with the test takers about how they made decisions about the RSES scores. With respect to the validity evidence for the RSES scores, this focus means that the interpretations of the RSES scores will take into account test takers’ understandings and meanings of the test items and test scores (i.e., participants’ or experiential validity).

Explanatory-interpretative model. Kagan (2012) noted that "personal interpretations of an experience are basic in psychological theory as cells are in biology" (p. 336) and that
"interpretations of an event transcend the brain response to an event" (p. 337). These comments represent a timely reminder that no viable psychological theories are possible without taking into account subjective and inter-subjective interpretations of events or of item content. However, there are no consistent and effective interpretative models in testing and psychometrics and the default assumption is that there is one given meaning and interpretation for each construct targeted by tests. Hence, the proposed model attempts to incorporate hermeneutic-interpretive frameworks with the intention to complement the explanatory ones.

**Theoretical model.** I will start this section by providing a brief review of the main theoretical and empirical findings about the self-esteem construct measured by the RSES (*Table 2.1*). Then, given that the RSES is a self-report measure of self-esteem, I will discuss the theoretical or expected response processes associated with responding to self-report items (e.g., self-processes, and self-related cognitive and affective response processes). A theoretically informed model of responding to the RSES items will then be presented and further elaborated (*Figure 2.6*). This model will integrate theoretical and empirical knowledge about the self-esteem construct (presented in *Table 2.1*), and substantive knowledge from the studies on self-processes and neuro-phenomenology of self as it applies to responding to self-report items (presented in *Figure 2.5* and *Table 2.2*).

**The construct of self-esteem.** A conceptual review of the theoretical perspectives and empirical findings regarding the self-esteem construct (Zeiggl-Hill, 2014) is presented in *Table 2.1*, and its purpose is to highlight the theoretical elements that have been important in the conceptualization of this construct and their implications on elucidating the theoretically expected response processes underlying answering the RSES items.
Table 2.1
Theoretical Underpinnings of the Self-Esteem (SE) Construct

<table>
<thead>
<tr>
<th>Theoretical stance</th>
<th>Representatives</th>
<th>Main tenets</th>
<th>Implications for response process model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-evaluative construct</td>
<td>James (1890)</td>
<td>SE is a self-evaluative construct with respect to one’s competence and self-worthiness</td>
<td>Self-evaluative processes with respect to competence, self-liking and self-worthiness</td>
</tr>
<tr>
<td></td>
<td>Rosenberg (1965)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivational construct</td>
<td>Adler (1927)</td>
<td>SE is a dynamic construct that motivates people to maintain or restore a baseline of self-esteem; Need for self-esteem as basic human need</td>
<td>Motivational processes such as restoring, compensating or maintaining self-esteem</td>
</tr>
<tr>
<td></td>
<td>Maslow (1943)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rogers (1960)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attitudinal construct</td>
<td>Rosenberg (1965)</td>
<td>SE is a stable attitude towards one’s self characterized by content, direction, intensity and stability</td>
<td>Socially shaped processes in developing this stable attitude towards self</td>
</tr>
<tr>
<td>Evolutionary interpersonal construct</td>
<td>Leary (2004)</td>
<td>SE has evolved to ensure social belongingness and to regulate social interactions</td>
<td>Response processes related to social regulation of behaviour</td>
</tr>
<tr>
<td></td>
<td>Bucur (2007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Existential construct</td>
<td>Pyszczynski et al., 2004a</td>
<td>SE represents an existential buffer against death anxiety (Terror Management Theory)</td>
<td>Death salience= self-esteem threat/trigger</td>
</tr>
</tbody>
</table>

In sum, self-esteem can be conceptualized as a socially-shaped, interpersonal concept (Oyserman, 2004), evolutionarily designed to ensure belongingness (Leary, 2004), and to offer a buffer against existential anxiety (Pyszczynski et al., 2004a). Self-esteem is not only a static self-evaluative construct (James, 1890; Rosenberg, 1965) but a dynamic, motivational one (Maslow, 1943; Rogers, 1980), in the sense that individuals are motivated to maintain a certain level of self-esteem, and, thus, engage in actions and behaviours meant to ensure the homeostasis of their
perceived self-esteem (Branden, 1994). Self-esteem also represents a stable attitude towards self, shaped by socio-cultural values and expectations, and can be described in terms of content, direction/valence, intensity and stability (Rosenberg, 1965). Competence and self-liking represent the main aspects of the self-esteem concept (James, 1890; Rosenberg, 1965; Zeiggler-Hill, 2014). The sources of self-esteem can be internal (e.g., self-beliefs, values, expectations) or external (e.g., feedback from others, comparisons with others), and, thus, in reporting self-evaluations, people access content from these sources (Epstein, 1995).

Therefore, we can expect that during answering the RSES items, respondents will engage in response processes consistent with these theoretical dimensions: self-evaluative processes, and motivational-dynamic processes meant to protect or maintain self-esteem by accessing relevant sources for the self-esteem evaluations in the context of socially relevant situations (e.g., failure, social exclusion). Given the social and evolutionary purpose of the self-esteem construct (i.e., ensuring social belongingness), we can hypothesize that social situations may serve as triggers for self-evaluative moments with respect to one’s self-esteem (i.e., ‘self-esteem moments’) as well as for the restorative processes.

**Self-reported self-esteem.** The RSES is a traditional self-report measure that, by design, is limited in terms of the information that it yields, and with respect to the response processes engaged by respondents during answering these items. Using Berkovich-Ohana and Glicksohn’s (2004) terms, self-reported self-esteem represents a product of the “second order access awareness” (p. 7) in contrast with the first order phenomenal awareness or in the moment awareness. The second order awareness is a reflective awareness that is explicit, conceptual and accompanied by focused attention to internally generated input (Gallagher & Zahavi, 2008). This type of awareness recruits brain areas (i.e., the fronto-parietal network) that make possible the
intentionality of the decision making processes involved in rating test items, and the verbal, language mediated self-reports as well as the possibility of socially desirable responding.

Therefore, when answering self-report items such as the RSES items, test takers are expected to focus their attention inwardly on relevant self-contents and experiences (i.e., self-processes), and to engage in intentional and linguistic communication (i.e., verbal self-reports) about relevant self-contents (i.e., declarative self-knowledge) or personal memories and experiences accompanied by self-related cognitive processing (e.g., autobiographical memory) and self-related emotionality (e.g., embarrassment, pride). It is then expected that the following response processes will underlie answering self-report items: a) self-processes, b) self-related cognitive processes, and c) self-conscious emotional processing. The next sections will discuss in more detail each of these response processes.

**Self-processes.** Self-related processes or self-processes are those response processes that are mobilized when someone reflects and reports on characteristics of one’s self or self-constructs such as: self-concept, self-image, self-efficacy or self-esteem. Research studies have identified that self-processes possess a neural signature that distinguishes them from the supportive cognitive and affective processes also engaged when people reflect upon characteristics related to self (Zahavi, 2003). The main categories of self-processes identified in the research literature and presented in *Figure 2.5* are: (1) the self-referential processes involved in reflecting and reporting about self, using 1st person references, and developing self-focused temporal, causal narratives, (2) the self-evaluative processes involved in evaluating and monitoring important self-characteristics, (3) the self-regulation processes meant to regulate one’s sense of self and self-characteristics in the ongoing interaction with the world, (4) the self-attribution processes involved in deciding the extent to which an external circumstance is related
to self, and (5) the self-expression processes meant to ensure the proper outer expression of self-characteristics.

![Figure 2.5. Self-Processes Underlying Responding to Self-Report Items](image)

Self-esteem represents both an outcome of self-processes (e.g., self-esteem is the result of self-evaluation processes), and also an intrinsically dynamic construct (i.e., the need for self-esteem) capable to ‘turn on’ a chain of self-processes meant to build, maintain, protect, restore, or enhance the perceived level of self-esteem (i.e., the self-regulation processes). Hence, I expect that the self-processes presented in Figure 2.5, particularly the self-evaluative, and self-regulatory processes (Lewis, 1995; 2000; Tangney & Dearing, 2002) will play a central role in the model of response processes underlying answering the RSES items.
So far, of the five self-processes presented in Figure 2.5, only the self-referential processes have been empirically investigated in relationship with answering self-report items. Hsu et al. (1989) and Holtgraves (2004) conducted research studies about self-referenced responding when answering self-report measures in an honest versus fake manner. They noted that the process of faking during self-reporting engaged very limited self-referenced responding (e.g., accessing declarative knowledge rather than autobiographical memories) and respondents tended to provide quick and minimally elaborated answers. More specifically, Hsu et al. stated that “whereas honest responding leads to a self-referenced interpretation of the item content, faking leads to a purely semantic item interpretation” (p. 29). For example, instead of self-referenced processing, respondents assigned to the fake experimental condition employed direct retrieval and semantic item interpretation in order to answer self-report questions. This propensity was confirmed by eye tracking measures that suggested a quick and early fixation on the semantically correct response with a tendency to choose extreme responses in the desirable direction, rather than choosing neutral or moderate ratings based on engaging in self-referential processes.

These findings are relevant for understanding the theoretical response processes model of self-esteem because they help distinguish phenomenologically the self-referenced or self-referential processes from the semantic, cognitive, and quick processing involved in fake self-reported responses. Specifically, an exclusive use of semantic responding during answering the RSES items would be regarded as a possible indicator of superficial engagement with items or possible fake responding.
Self-processes and response strategies. Elaborating on the self-processes illustrated in Figure 2.5, Table 2.2 presents the associated response strategies expected to be observed during responding to the RSES items.

Table 2.2

Self-Processes and Response Strategies Involved in Answering the RSES Items

<table>
<thead>
<tr>
<th>Category</th>
<th>Triggers</th>
<th>Response strategies</th>
<th>Response processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-referential</td>
<td>Self-reflective activity/task</td>
<td>Accessing and retrieving autobiographical memories; Spontaneous linking self to others; decentering and intersubjectivity; Heightened accessibility to elaborated contents about self (traits, implicit theories, beliefs about self, self-narratives)</td>
<td>Explicit use of first person reporting Spontaneous self-others comparisons Inferred self-referential reporting Temporal processes</td>
</tr>
<tr>
<td></td>
<td>Introspection</td>
<td>Causal, temporal self-narratives</td>
<td></td>
</tr>
<tr>
<td>Self-evaluation</td>
<td>Situation of threat to self and self-esteem</td>
<td>Accessing objective information about self (e.g., test results, achievements, successful relationships)</td>
<td>Self-assessment</td>
</tr>
<tr>
<td></td>
<td>Past (evoked) situations of rejection or failure</td>
<td>Temporal comparisons (with past and future selves)</td>
<td>Self-verification</td>
</tr>
<tr>
<td></td>
<td>Anticipated situations of failure or rejection</td>
<td>Social comparisons (upward and downward)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anticipation of a challenging situations</td>
<td>Lateral social comparisons with peers seen at the same level as the respondent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A very significant, valuable situation in the present or future</td>
<td>Asking for feedback (e.g., asking a friend or a trusted person for positive feedback)</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Triggers</td>
<td>Response strategies</td>
<td>Response processes</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>Actual or perceived threat to self-integrity and self-esteem</td>
<td>Ignoring negative feedback</td>
<td>Maintaining</td>
</tr>
<tr>
<td></td>
<td>Actual or perceived threat to self-coherence and consistency</td>
<td>Buffering (e.g., distancing, withdrawing)</td>
<td>Restoring</td>
</tr>
<tr>
<td></td>
<td>Overwhelming situation for self that poses threat to self-integrity</td>
<td>Searching and drawing on sources of positive feedback</td>
<td>Compensating</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Enhancing</td>
</tr>
<tr>
<td>Self-expression</td>
<td>Situations that require exposure of self (e.g., social situations with high self-evaluative risk)</td>
<td>Minimizing the negative</td>
<td>Social desirable responding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximizing the secondary benefits of a certain situation</td>
<td>Impression management</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fake responding (e.g., overreliance on semantic processing)</td>
</tr>
<tr>
<td>Self-attribution</td>
<td>Situation or stimulus potentially relevant for self, either congruent or incongruent with self-values/goals/expectations</td>
<td>Assessing the motivational relevance (self-appraisal)</td>
<td>Stable, personal and global attributions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evaluating the motivational congruence</td>
<td>Temporary, situational and specific attributions</td>
</tr>
</tbody>
</table>

**The cognitive-affective processes.** Self-esteem represents a motivational-evaluative structure that enables a person to assess oneself in terms of personal worthiness and competence.

According to Damasio (2000, 2003a), Schachter and Singer (1962), and Lambie and Marcel (2002), any conscious self-evaluative structure is supported by processes of cognitive appraisals (attributions), and autobiographical memory, as well as valence (affective) conceptualization.

Hence, it is important to include the cognitive-affective processes among the core processes that comprise a theoretical response process model of responding to the RSES items, in addition to the self-processes presented in Figure 2.5.
**Self-related cognitive processes.** In order to generate the self-evaluations required by responding to the RSES items, several cognitive processes have to be activated, such as the episodic, autobiographical memory (Conway, 2009; Russel & Davies, 2012; Tulving, 1985; 2001), the declarative self-knowledge (i.e., the entire system of the self-representations), and the socio-cognitive mechanisms that support the self-attribution processes (Tyler, Cramer & John, 2009).

**Self-related affective processes.** As an evaluative construct, self-esteem is closely connected with basic emotions in general (e.g., sadness, fear, anger), along with the main self-conscious emotions, such as shame and pride (Tracy & Robins, 2004), as well as with core self-affect and valence/reflective affect (Berkovich-Ohana & Glicksohn, 2014). According to research findings in the neurophenomenology of affective experience (Lambie & Marcel, 2002), any evaluative processes such as those mobilized during answering self-esteem items engage self-related affective processes.

Research studies have distinguished between (a) the core affect representing the embodied, phenomenal, ‘in the moment’, un-reflected emotionality, and (b) valence affect representing the reflective, conceptualized, attitudinal, and evaluative type of affect (Berkovich-Ohana & Glicksohn, 2004). On an explicit self-report measure of self-esteem like the RSES, it would be reasonable to observe predominantly reflective, attitudinal, value-based self-evaluations and associated valence affect as well as the activation of self-conscious emotions (i.e., pride or shame). Thus, these affective processes will be further included in the response process model of self-esteem.

**Proposed theoretical response processes model.** In spite of the many research studies about self-esteem construct, currently no theoretical model of self-esteem has been proposed
(Zeiggler-Hill, 2014). Hence, the theoretical model proposed in this dissertation (Figure 2.6) represents a stand-alone theoretical contribution to this field.

After scrutinizing Rosenberg’s self-esteem theory (Rosenberg, 1965), and the relevant empirical research findings on the RSES (Zeiggler-Hill, 2014), it seems that, theoretically, the RSES targets or measures a global construct of self-esteem broadly defined as a positive attitude towards one’s self, which stems from a series of self-evaluation processes with respect to several key self-characteristics/aspects, such as: a perceived sense of competence or mastery, and a perceived sense of self-worthiness or personal value. The results of these self-evaluation processes set in motion a series of self-regulatory processes (restore, protect, defend) which eventually determine the content, stability (stable or labile), direction (positive or negative), and intensity (degree from low to high) of the reported self-esteem, and, implicitly, of the ratings on the RSES items. Hence, it is expected that the RSES scores and their descriptive qualities (i.e., content, direction, intensity and stability), are the product of the underlying mechanisms of self-evaluation and self-regulation meant to maintain and enhance the ‘good self’. In order to complete these self-evaluations, test takers can access several sources (e.g., personal events, social comparisons, relationships, and feedback from others), and, hence, report on these when answering the RSES items (e.g., the content of test takers’ responses). These construct specific processes are supported by more general self-processes (e.g., self-referential responding), and self-related cognitive and affective processes (e.g., autobiographical memory, self-conscious emotionality).

In attempt to bring these elements together, Table 2.3 presents the main response processes underlying responding to the RSES items, and Figure 2.6 illustrates the theoretical model of response processes underlying answering the RSES items. These response processes
will be further included in the theoretically informed coding scheme that will be used in the empirical evaluation of the model in chapter 3 of this dissertation.

*Table 2.3*

**Response Processes Underlying Responding to the RSES Items**

<table>
<thead>
<tr>
<th>Category</th>
<th>Response Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-processes</td>
<td>Self-referenced responding</td>
</tr>
<tr>
<td></td>
<td>Self-evaluation</td>
</tr>
<tr>
<td></td>
<td>Self-regulation</td>
</tr>
<tr>
<td></td>
<td>Self-attribution</td>
</tr>
<tr>
<td></td>
<td>Self-expression</td>
</tr>
<tr>
<td>Self-related affective</td>
<td>Valence conceptualization</td>
</tr>
<tr>
<td>processes</td>
<td>Self-conscious emotions</td>
</tr>
<tr>
<td>Self-related cognitive</td>
<td>Autobiographical memory</td>
</tr>
<tr>
<td>processes</td>
<td>(narratives of personal events)</td>
</tr>
<tr>
<td></td>
<td>Declarative self-knowledge</td>
</tr>
<tr>
<td></td>
<td>(beliefs, traits)</td>
</tr>
</tbody>
</table>
Figure 2.6. Theoretical Response Processes Model of Answering the RSES Items
Situated or contextualized model. Rosenberg (1965) explicitly recognized that self-esteem is significantly shaped by a variety of social and cultural processes. In addition, Leary (2004) and Bucur (2007) highlighted the interpersonal evolutionary aspect of the self-esteem construct. Therefore, self-esteem can be considered a socio-culturally situated construct activated in interpersonal situations. Specifically, self-esteem is viewed as a “social product” spawning from a combination of “one’s own internal perceptions and the external feedback received through interpersonal relationships” (Oyserman, 2004, p. 11). Considering that an individual desires to be accepted by others, it is understandable that one takes into account the perceptions of others when determining an attitude toward the self. Therefore, it is expected that interpersonal, social, and cultural processes will have a significant impact on the test scores obtained from a self-report measure of self-esteem. Moreover, given the interpersonal nature of the testing situation, it is likely that some of the elements of that situation may have an impact on the self-reported scores on the self-esteem measure.

For example, Ardilla (2005) summarized the main contextual, culture dependent elements that underlie a testing situation and shape test takers’ responses:

(1) One-to-one relationship, (2) Background authority, (3) Best performance, (4) Isolated environment (5) Special type of communication, (6) Speed, (7) Internal or subjective issues, and (8) The use of specific testing elements and strategies. In addition, it is proposed that ‘the distance’ (e.g., gender, age, ethnicity) between the examiner and the examinee may potentially impact the testing situation (p. 185).

Each of these variables interacts with the proposed process based model of self-esteem and may affect the reported ratings on the self-esteem measure. The “person X situation” model suggests that people mobilize prototypical responses that may vary across situations (Mischel &
Shoda, 1995). No research studies have investigated this for the RSES. However, I hypothesize that at least several of these situational characteristics may play a role in responding to the RSES using TAP and CI, specifically: the one to one relationship (association dimension), background authority (dominance dimension), isolated environment, and subjective states (intimacy-formality dimension).

These interpersonal dimensions are all potential triggers for the self-esteem processes described in the previous section of this chapter and, thus, are very relevant to the situated process based model. For example, self-esteem research has emphasized the need to maintain and restore self-esteem when faced with situations that are potentially threatening, such as social situations involving rejection, exclusion, negative self-evaluations, or failure. Social situations are the breeding ground for activating self-esteem processes such as enhancement, maintaining, and verifying. Thus, the testing situation may be already activating for an assessment of self-esteem. For instance, the authority dimension present in the testing situation may easily trigger memories of injured self-esteem via external negative evaluation, and being asked to share personal content (i.e., the intimacy dimension) may lead to feeling vulnerable to rejection, another trigger for self-esteem processes. It is timely to note that, although some situated or contextual response processes may be theoretically anticipated, many of them could also emerge at the intersection between test taker and the testing situation so they cannot be fully anticipated from the beginning. A visual illustration of the socially situated, contextualized model of response processes underlying responding to RSES items is presented in Figure 2.7.
Figure 2.7. Situated Response Processes Model

TEST ITEMS

SOCIOCULTURAL VALUES

SOCIAL CONTEXT

TEST ADMINISTRATOR

P E R S O N

THEORETICAL MODEL OF RESPONSE PROCESSES

TEST SCORES
**Integrative Model.** The theoretically expected and contextualized response processes fit well with the interpretive, teleological and socio-cultural contextualized epistemological commitments that are at the foundation of this model. Given the social, interpersonal nature of the self-esteem construct itself as well as the influence of culture on the self-esteem (Kitayama, Markus & Matsumoto, 1995), it becomes necessary to adopt a contextualized and situated epistemology to explain or interpret the scores on the RSES.

Adopting a socio-culturally situated interactionist world-self framework, we can imagine the person/self in continuous dialogue with the world as a set of socio-cultural meanings, and interpersonal situations, including good experiences (e.g., belonging, acceptance, success, achievements) and unpleasant experiences (rejection, failure, exclusion, death, losses). Even the testing situation is structured as a social situation alongside affiliative, authority and intimacy dimensions which may evoke or create the opportunity for experiencing ‘self-esteem moments’. The self pole is comprised of internalized interpersonal contents (e.g., self-representations, self-beliefs, self-narratives, self-models, self-related or activated emotions) that can be accessed via autobiographical memory and the associated self-processes (e.g., awareness, evaluations, memory). Then, we can hypothesize the unfolding of the self-esteem response processes model as follows:

1. The “I“, embedded in a socio-cultural context, is aware of a situation arising and of the fact that it pertains to self (e.g., “it is about me”). For example, the testing situation is a social situation that contains a few powerful triggers for self-esteem and certain test items may also evoke these situations. The self-awareness processes are engaged at this stage and signal that something involving the self is happening, and it could impact or even threaten the positive view of self.
(2) The “I” attempts to assess the extent and nature of the self-involvement by initiating a series of self-appraisal and self-attribution processes to assess the motivational relevance, and the motivational congruence of the given situation. Also, existent and previous self-esteem levels are assessed and used as baseline to assess whether self-esteem is lowered or heightened by the event. To answer how relevant the triggering situation is for self, one needs to access self-relevant domains connected with values endorsed by self (e.g., what is important for that person or what does the person value?), wishes and expectations, all these shaped by social meanings and cultural factors.

(3) If the result of this appraisal process is that the situation is relevant for self, and it potentially presents a threat to the positive view of self (self-esteem), then the next question is: “is this my fault or my responsibility?” Answering this question mobilizes the self-attribution processes. If the attributions made are internal and stable, then self-conscious emotions are likely to be experienced (e.g., shame and guilt), and these emotions further signal a threat to self-esteem which eventually mobilizes self-evaluation and self-regulation processes.

(4) Depending on the result of the self-evaluation processes, the self-regulation processes prompt behaviours meant to restore, protect, maintain or enhance self-esteem, and they may take the shape of social and temporal comparisons (e.g., distancing from bad self, retrieving compensatory autobiographical events), coping by trying to modify the external situation, asking for self-esteem restorative feedback, seeing self in a better light than it really is (self-serving bias), projecting and blaming, rationalizing, impression management, as well as overt behaviours such as asking for feedback, trying to impress, lying etc. Self-regulation processes may lead to alteration of the presented self-concept, impression management etc.
(5) This dynamic of the model is shaped by the relevant socio-cultural meanings, and the specific of the situation or context in which this process is taking place, including teleological motivations such as test takers’ hopes, wishes or desires.

**Conclusion.** The model of the response processes underlying answering the RSES items consists of four pillars: a) epistemological foundations grounded in socio-cultural and epistemic contextualism, and teleological-interpretative stances, b) theoretical underpinnings integrating construct theory, self-processes (i.e., self-evaluation, self-regulation), and self-related cognitive and affective processes, meant to provide and maintain an evaluative view on self in terms of competence and self-liking, c) situated-contextualized dimensions such as: interactive interpersonal processes, social meanings, linguistic dimensions, and cultural factors, and d) integrative model framework bringing together the epistemological commitments, theoretical grounds and contextualized dimensions in order to propose an explanatory-interpretative account of the RSES scores. Chapter 3 of this dissertation will present an empirical evaluation of this model of response processes associated with responding to the RSES items.
Chapter 3: Examining the Substantive Validity of the Rosenberg Self-Esteem Scale Using Response Processes Inquiry

Introduction

This chapter presents an empirical investigation of the response processes that are involved in answering self-esteem self-report items (i.e., the Rosenberg Self-Esteem Scale) using the theoretical model of response processes developed in the previous chapter of this dissertation. The chapter begins by stating the purposes and research questions of this study, and continues by providing a description of the participants, procedures, and measures used in this study. Next, data analysis strategies, and the study findings will be presented. The results of this study will be discussed in light of the theoretical model of response processes underlying answering the RSES items. A discussion about the contributions of the study and future research directions will conclude this chapter.

Purposes and Research Questions

The Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) is a well-known and widely used measure of self-esteem in social science research and clinical practice, and yet no study has examined the substantive validity of this scale using response processes inquiry as described by the Standards (2014), although Gray-Little, Williams and Hancock (1997) investigated the response patterns to RSES items using IRT methodology. Thus, one of the purposes of this research study was to investigate respondents’ mental processes when they completed the RSES. Another equally important purpose of this investigation was to empirically evaluate and possibly revise the proposed theoretical model of response processes underlying answering the RSES items (Figure 2.6 and Figure 2.7).

Specifically, this study answered the following research questions: (a) what response processes did test takers demonstrate while answering the RSES items, (b) to what extent were
respondents' observed response processes while answering the RSES items consistent with the anticipated response processes suggested by the theoretical response process model, and (c) what is the contribution of response processes inquiry to the substantive validity of the inferences made from the RSES scores?

**Method**

**Participants.** The sample of this study consisted of 30 adult participants selected from the general adult population of British Columbia. It is noteworthy that the average sample size used in studies examining response processes was around 20-25 participants, with some studies reporting a sample size of 16 whereas another study reported a sample size of 67 (Padilla & Benez, 2014). Thus, the sample size of 30 for the present study falls within the typical range for this type of investigation.

In terms of the sample composition, 18 women and 12 men participated in this research. The average age of the sample was 41, with the youngest participant being 20 and the oldest 56. Fifteen participants identified themselves as Caucasian, ten identified themselves as South-East Asian (i.e., Phillipina and Vietnamese), and five described themselves as mainland Chinese. Most participants (n=26) reported education levels at university degree level and above, 2 participants endorsed high-school diploma as their highest level of education, and 2 participants mentioned that they completed post-secondary trade training.

**Exclusionary criteria.** Individuals who could not understand and communicate fluently in both oral and written English were excluded from the study given the amount of verbal self-reflection that participants needed to do during this research project. In addition, persons with a self-reported diagnosis of mental illness were excluded from the study given that this study involved remembering and reflecting upon potentially triggering experiences of social rejection,
failure or disappointment, during answering test items and while taking part in cognitive interviewing. These characteristics of the study could have potentially exacerbated the emotional distress of people who were already experiencing an emotional disturbance, and, thus, these potential participants were screened out. Lack of self-reported previous mental health diagnosis was not a guarantee that some participants were not emotionally distressed by the study; thus, at the end of the session, I provided all participants with a short debriefing interview about their experiences during the study and a list of resources that they could access in case they needed to further process any feelings or concerns triggered by the present study.

In order to screen out prospective participants meeting the exclusionary criteria described above, I talked to them by phone before making an appointment for participating in the study, and I asked them about previous mental health history and/or diagnosis and about their ability to communicate in English. The latter ability was also informally evaluated during the conversation that prospective participants were able to carry out over this initial phone contact.

**Recruitment.** Research participants were recruited through posters (e.g., placed in different locations such as: colleges/universities, community or recreation centres, churches), in-person announcements to classes and community centres, e-mail announcements, Craig’s list, Facebook, and snowball sampling (i.e., word of mouth notification of the study by the researcher or study participants). The recruitment poster used for this study is provided in Appendix A. At the beginning of the research study, all participants were informed about the purpose of the research, conditions, and any financial compensation, and signed an informed consent form (Appendix B). They also completed a Demographic Sheet (Appendix C). Ethical approval for the study was obtained from the UBC Behavioural Research Ethics Board.
**Procedure.** This study used cognitive interviewing during the completion of the Rosenberg Self-Esteem Scale. In essence, participants were asked to first think aloud and then reflect upon the researcher’s probes as they completed this self-report measure. Details about the measures used in this study are provided under the Measures heading.

**Cognitive interviewing.** Cognitive interviewing is a broad class of methods that aim at eliciting verbal information from respondents while they answer self-report or survey items, with the purpose of achieving a better understanding of the mental processes employed by respondents when answering items (Willis, 2005). Within the broader category of cognitive interviewing, two main paradigms have emerged and become differentiated: think aloud protocols (TAP) and verbal probing (VP), although a blend of both methods has been used in research studies.

*Think Aloud Protocols* (TAP; Ericsson & Simon, 1980) represent a class of methods that ask participants to think aloud as they are answering items on a questionnaire or when they are engaged in a task. Respondents are asked to verbalize whatever they are looking at, thinking, doing, and feeling as they go about completing or recording their responses (i.e., concurrent verbalization). This method allows one to study respondents’ processes rather than just their final answers on a self-report measure.

*Verbal probing* (VP) consists of asking questions or probes about respondents’ cognitive, motivational, and emotional processes as they complete self-report measures. The model that I used for verbal probing is the Survey Interaction Model that takes into account not only the cognitive processes involved in answering items but also respondents’ emotions, motivations, and attitudes when they answer these items (Jobe, 2003). As a result of using more sensitive
questions or probes, this model allowed me to interpret the shifts that could possibly occur in participants’ answers.

In the present study, I integrated TAP with VP, whereby, after being encouraged to think aloud freely, respondents were probed for specific information by the interviewer, such as asking the respondent to reformulate an item or define some of the key terms in their own words. While TAP allowed for more spontaneous and in the moment track of respondents’ mental processes, VP allowed for a more in-depth and targeted exploration of participants’ cognitive and conative processes of interest. Appendix D provides examples of the probing questions used during CI.

**Study timeline.** First, all participants were explained the purpose, format, and procedure of the study, and were offered the opportunity to ask for clarifications regarding any of these aspects before signing the informed consent form. After providing informed consent, all participants completed the baseline questionnaires in the same order: the single item self-esteem measure, and then the Rosenberg Self-Esteem Scale, while engaging in cognitive interviewing. Throughout the completion of the self-report items, participants were encouraged to reflect upon how they chose their responses and what mental processes were mobilized while completing the task. Respondents’ TAP and their responses to verbal probes were digitally audio-recorded and were later transcribed and analyzed. At the end of the entire study session, participants completed a demographic sheet. Participants were also offered a package of counselling resources that they could access in the case that they needed further support (Appendix E).

**Measures.** All participants completed the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) and a demographic questionnaire. Only the ratings on the RSES were examined using the TAP and cognitive interviewing.
Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). The RSES is a widely-used self-esteem measure in social science research. It consists of 10 items that are answered on a four point Likert-type scale ranging from ‘strongly agree’ to ‘strongly disagree’ (Appendix F). Five of the scale items are positively worded statements and five are negatively worded ones. The scale is believed to measure state self-esteem by asking the respondents to reflect on their current feelings. The RSES is considered a reliable and valid quantitative tool for self-esteem assessment (Blascovich & Tamaka, 1993).

Demographic Questionnaire. The purpose of the demographic questionnaire was to provide descriptive information about participants. Information was collected about the following variables: age, birth date, gender, level of education, and ethnic/racial/cultural background (Appendix C).

Data Analysis

Data checking and transcribing. First, data were checked for accuracy, and then, the audio-recordings of the TAP and cognitive interviewing sessions were transcribed verbatim by a professional transcriptionist. After all 30 transcriptions were completed, I listened to each tape while reading the corresponding transcript so that I could check the transcripts for accuracy and to achieve a thorough understanding of the information that was further analyzed.

Content analysis. The 30 transcripts were analyzed with the aim of understanding the mental processes involved in responding to the RSES items in order to gather substantive validity evidence for the interpretations made of the RSES scores. Content analysis unfolded in three steps: (1) theoretically informed coding, (2) inductive, open coding, and (3) examining the response patterns on the RSES.
**Theoretically informed coding.** Using the framework of the response processes model elaborated in chapter 2, a theoretically informed coding scheme was developed and implemented during this phase of data analysis (Appendix G). A qualitative item level analysis was performed using the theoretically established coding categories in order to identify the response processes and contents associated with participants’ responses to the RSES items. First, each transcript was read in order to get a global, overall understanding of the participants’ narratives, and to identify any unusual ways of responding to any of the RSES items. Then, the responses to each item were coded using the theoretically established codes and coding categories presented in Appendix G until these codes were exhausted. The detailed results of this phase of data analysis are presented in Appendix H.

**Inductive coding.** During this step of data analysis, the coded transcripts were reviewed with the purpose of discerning contents and themes that had not been included in the theoretically informed coding scheme. I used open coding to code for the content that was not accounted for by the initial theoretical coding. These response processes and contents that were not included in the initial coding scheme but were recognized in the data were coded separately, and are presented in Table 3.1.

**Examination of response patterns.** An examination of the response patterns on the RSES was conducted with the purpose of identifying how participants made use of the 4-point Likert-type response scale for each item, and the overall pattern of responses to the RSES items. In addition, throughout the data analysis process the researcher kept track of any relevant information with respect to unusual or idiosyncratic ways of responding or interpreting the meaning of the RSES items, and this information was included in the Results section of this chapter.
Results

For a detailed presentation of the results of the item level analysis for each of the 10 items of the RSES the reader is encouraged to refer to Appendix H. For the sake of brevity, the next paragraphs will summarize the main global trends in responding to the RSES items in terms of: (a) response processes, (b) patterns of item endorsement, and (c) participants’ understanding of the RSES items.

Response processes. This section will present the main categories of response processes that were observed while completing the RSES as well as some of the strategies of responding to these items (e.g., semantic versus self-referenced responding).

Self-evaluation and self-regulation. Self-evaluation and self-regulation were the most frequently observed response processes underlying answering the RSES items, and they seemed to be differentially involved across various items. Whereas positively worded items (items 1, 3, 4, 7 and 10) engaged mainly participants’ self-evaluative response processes (e.g., self-improvement, self-assessment, and self-verification via social and temporal comparisons) in light of their stable, trait-like personality characteristics, the negatively worded items (items 2, 5, 6, 8 and 9) activated primarily respondents’ self-regulation processes (e.g. restoring, compensating, normalizing or maintaining self-esteem levels) in conjunction with the relevant self-evaluation processes. In particular, the agreement with these latter items (i.e., responding with “agree” or “strongly agree”) led to a plethora of self-regulation strategies meant to protect or enhance one’s self-evaluations.

Self-critical versus reflective-agentic evaluation. Items 2, 5, 6, and 9 mobilized predominantly participants’ self-critical evaluation processes (i.e., measuring oneself against others’ or own expectations, falling short of these expectations, or struggling to meet expectations). Items 1, 3, 8 and 10 predominantly engaged agentic self-evaluation processes in
terms of moral values, personal integrity, and self-transcendence. It is noteworthy that most of the items that predisposed participants to engage in reflective-agentic self-evaluation processes were the same items that solicited anticipatory-teleological processes, which suggest that they may be part of a common more generic response process or mechanism.

**Cognitive/semantic and affective/self-referential processing.** Responding to negatively worded items (i.e., items 2, 5, 6, 8, and 9) predisposed participants to experience and report a higher amount of self-conscious emotionality (e.g., shame or guilt) stemming from remembering or re-experiencing events related to failure, interpersonal criticism, exclusion, or rejection, which represent the typical triggers for evaluating and regulating self-esteem. These items appeared to predispose participants to engage primarily in processes of autobiographical recall (e.g., experiences of failure and interpersonal exclusion or rejection) and self-referenced responding during which respondents accessed primarily personal, self-relevant contents. These items also generated the most polarized or extreme response tendencies, and it may be that affective processing and autobiographical retrievals led to a faster and clearer endorsement of the items, either positively or negatively.

In contrast, cognitive processing was predominant in responding to the positively worded items (items 1, 3, 4, 7 and 10), and this was consistent with the primarily semantic processing observed during responding to these items. Cognitive processing involved self-appraisal and self-attribution processes, declarative memory, and counterfactual thinking. In addition, items 1, 3, 4 predisposed participants to engage with the item content in a predominantly semantic fashion. Specifically, in coming up with answers to these items, respondents accessed primarily declarative self-knowledge (e.g., beliefs about self, implicit theories about self), and, much more infrequently, autobiographical memories. For instance, item
1, due to its general content (“On the whole, I am satisfied with myself”) solicited very heterogeneous response contents, and respondents’ engagement with the item was mainly done through remembering their social roles and accomplishments. Similarly, while responding to items 3 and 7, participants mobilized primarily declarative self-knowledge (e.g., stable and broad self-descriptions, and implicit theories about self and others), and semantic processing. The expression “at least on an equal basis with others” (item 7) made it easier for participants to answer the item by prescribing the response strategy (i.e., social comparisons) but predisposed them to respond in terms of social roles and general beliefs about self. In addition, responses to item 10 appeared significantly shaped by the Western cultural messages and pressures to be or try to be positive, and, thus, this item elicited lots of stable self-attributions, and socially prescribed or expected responses.

**Socially desirable responding.** Some participants spontaneously mentioned during the cognitive interviewing that they engaged in a socially desirable responding style when answering items 2, 5, 6 or 9. In this sense, one participant told the interviewer: “I don’t want to tell you about these [not good at all] times. I would feel worse if I did”.

**Anticipatory and teleological response processes.** Items 1, 2, 3, and 8 solicited participants’ anticipatory and teleological response processes to a much higher degree than the rest of the items. In order to decide on their final score for these items, participants referred primarily to characteristics of their future self, possibilities and hopes relevant to self, as well as their desires and wishes. Constructing or imagining these future possibilities about self appeared to be one of the main strategies or mechanisms underlying responding to these items.

**Response processes not anticipated in the theoretical model.** Some of the response processes observed when responding to the RSES items were not included in the initial
theoretical model, and were noticed during the data analysis phase that followed the theoretically
informed coding. These response processes are presented in Table 3.1.

Table 3.1

Response Processes not Anticipated in the Theoretical Model

<table>
<thead>
<tr>
<th>Response Process</th>
<th>Response Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflective-agentic evaluative processes</td>
<td>Reflecting upon one’s values and general human values</td>
</tr>
<tr>
<td></td>
<td>Making choices/responding in light of these values not only in terms of meeting expectations like during the self-critical evaluation</td>
</tr>
<tr>
<td>Self-transcendence processes</td>
<td>Contributing to something or someone beyond self</td>
</tr>
<tr>
<td></td>
<td>Making a difference beyond one’s self-interests</td>
</tr>
<tr>
<td></td>
<td>Harmonizing self-interests with communal interests</td>
</tr>
<tr>
<td>Additional self-regulation processes</td>
<td>Normalizing</td>
</tr>
<tr>
<td></td>
<td>Self-talk</td>
</tr>
<tr>
<td></td>
<td>Reshaping/recalibrating the meaning of some items (e.g., “not no good but not good enough”; “not useless but helpless”)</td>
</tr>
<tr>
<td></td>
<td>Changing the response format for the Likert scale</td>
</tr>
<tr>
<td>Cognitive processes</td>
<td>Counterfactual thinking²</td>
</tr>
<tr>
<td>Embodied responses</td>
<td>“Gut feeling”</td>
</tr>
<tr>
<td>Interpersonal, dialogical processes</td>
<td>Struggling to come up with an answer and asking the interviewer for validation; taking items very seriously, dialoguing back and forth to decide on a final answer</td>
</tr>
</tbody>
</table>

² Counterfactual thinking represents the human tendency to create possible alternatives to life events that have already occurred; something that is contrary to what actually happened.
**Patterns of item endorsement.** This section will present the most commonly observed patterns of item endorsement in relation to the proposed 4-point Likert-type response scale used for responding to the RSES items.

**Extreme versus moderate endorsement.** To answer some of the RSES items, participants used almost exclusively the middle points of the Likert-type response scale (i.e., “agree” and “disagree”), whereas for other items participants tended to choose more extreme response categories (i.e., “strongly disagree” and “strongly agree”). Most of these tendencies seemed to be connected with specific qualifiers or to the item wording and content.

Items 1, 4 and 7 invited an almost exclusive use of the moderate categories “disagree” and “agree”. Specifically, only one of the 30 participants answered “strongly agree” to item 1 whereas all the others respondents chose either “disagree” or “agree”. This tendency to endorse moderate ratings seemed to be due to a large extent to participants’ anticipated possibilities and hopes that self could be improved in the future. In addition, for item 1, 16 out of 30 participants created a 2.5 category of the Likert scale named “between disagree and agree” in an attempt to illustrate their desire to improve in the future and the impossibility to decide more firmly at the time of the assessment. This response tendency might have been influenced by some sociocultural messages that tend to promote the belief that that people should always strive for the better (e.g., the better car, the better mobile phone, or the better partner), and that the future will be or could be better than the present.

Item 4 received the most critical comments about the lack of context, and the need to be more specific or contextualized. Subsequently, this item was the top item for spontaneous intermediate ratings (i.e., 2.5 meaning “I cannot decide because of lack of context” or “It depends so I prefer to be neutral”). Respondents described having a difficult time deciding
between “disagree” (2) and “agree” (3) on this item because of the lack of context and specificity. The Asian participants tended to contextualize this item the most, to the degree of adding new content and meaning to the item, and they expressed more desire for nuanced and more moderate ratings, mainly because of the hesitancy to put others down in downward social comparisons. Here are some comments from several participants describing their need for context and specificity in order to come up with the answers for item 4: “I think that ‘things’ is a little uh vague but we, of course we, everybody could do things”; “Depending on with whom you compare; I cannot decide, it depends”; “I guess it comes down to with whom do you compare yourself. I think always people are better than me but I’m always better than some, some other person”; “Depends on the context of what this means, right? So yeah, there’s a lot of areas where I’m not good at, you know, like I’m not a race driver, race-car driver or not a very good swimmer, you know, those kinds of things, right? So I guess it depends on what, what that means and then, there isn’t an answer on here that I can say neither agree nor disagree”; “I’m good in some areas but I’m also not good in other areas, right? So it depends. I’d say 2.5”

Item 3 and 7 were the items with the highest rate of positive endorsement (i.e., “agree” and “strongly agree”). None of the participants chose “disagree” or “strongly disagree” for these items. This preference appeared to be due to the qualifier “some” in item 3 and to the expression “on an equal basis with others” in item 7 that made respondents readily agree with the statement based on logical inferences (e.g., “everyone has some good qualities so I must have some too”).

Items 2, 5, 6, 8 and 9 were the items with the highest rate of more extreme endorsement both in the negative and positive direction. It is noteworthy that these items are the negatively worded items of the RSES, and contain some emotionally charged wording such as: “failure”, “not good at all”, or “not much to be proud of”. This tendency towards a faster and more extreme
responding may be due to some of this wording that has elicited autobiographical recall and the associated self-conscious emotionality that might have led to a more extreme style of responding based on affective processing.

**The role of qualifiers.** For some items, the qualifiers had a significant impact on the final ratings. For item 2, the qualifier “at times” seemed to exert a significant influence on how respondents made use of the Likert-type response scale. The qualifier “at times” appeared to prime the respondents to engage more often in temporal comparisons and in quantitative reasoning regarding the frequency, temporal stability, and duration of the “not good at all” moments. For these participants, the final score on this item depended on how often, how long, and how stable they deemed these moments to be. In addition, the qualifier “at times” tended to take precedence over “not good at all”: “At times I’m thinking I’m no good at all. Hmm. Interesting. Um, yeah. There are times. If I only have agree or disagree, I would think I would agree with that. Because it is “at times”. “At times” as qualifier also made people normalize the times when they felt “not good at all” and, in return, this increased the chance of an item to be endorsed (either with “agree” or “strongly agree”). In contrast, the few participants that emphasized “not good at all” in their decisions tended to choose extreme answers faster (“strongly disagree” or “strongly agree”), and ignored the qualifier “at times”.

The absence of appropriate qualifiers (e.g., item 4) generated participants’ critical comments about the lack of context and the need for the item to be more specific or contextualized. Subsequently, these items made participants come up with new intermediate ratings on the Likert-type response scale (i.e., 2.5 meaning “I cannot decide because of lack of context” or “It depends so I prefer to be neutral”).
Understanding of the RSES items. Some respondents either re-defined the meaning of some items or words or they changed the intended meaning of some of the RSES items. For item 6 ("I certainly feel useless at times"), some participants reframed the meaning of “useless” in relationship with “not good at all” from item 2. Some argued that the two terms meant the same thing whereas other participants decided to adopt the following distinction when answering these items: “‘Not good at all ‘is a subjective judgment about self, is how you feel about yourself, or think about you. If you are critical then you feel ‘no good at all’ but it is not objective because it is how you think. ‘Useless’ is objective: not being able to do something. It is for real. Like if I think that I am not good to move this chair that is subjective, but if I really cannot do it then it is objective so I am useless because I cannot do it”.

When answering item 7, participants defined the meaning of “person of worth” in two different ways: the innate self-worth as essential to all human beings, and the self-worth coming from contributions and actualizing the human being potential, and from contributing to others and to society. Depending on the meaning attributed to the expression “person of worth”, the rating of this item varied. Namely, respondents who used the definition highlighting the innate worth of all human being were prone to more positive and unequivocal ratings than participants who defined self-worth in terms of concrete personal contributions to others or society.

One of the most striking observations with respect to respondents’ different understanding of the RSES items was the difference between how North-American and Asian participants defined the concept of self-respect (item 8), failure, positive outlook, and self-esteem. Metaphorically speaking, for North-American participants, self-respect meant going softer on oneself in the sense of promoting self-care and self-compassion whereas, for Asian participants, self-respect meant being harder on oneself. Also, for North-American respondents,
self-respect seemed to be focused on taking care of the needs of self whereas, for the Asian respondents, self-respect meant primarily respecting others and others’ needs.

The definition and meaning of “failure” also differed cross-culturally. For North-American participants, “not having” and “not doing” enough were the main criteria for being a failure. For Asian participants, failure was defined mainly as not being able to contribute, or not being able to do valuable things for community, including harmonizing and living peacefully with others.

Because of the North-American cultural messages and pressures to be positive, item 10 (“I take a positive attitude towards myself”) elicited lots of stable attributions and motivational attempts to change attitudes. For Asian participants, taking a positive attitude represented an active process of motivating oneself to persist usually through engaging in encouraging self-talk whereas, for North-American participants, a positive attitude was understood as the socially prescribed silver lining mentality linked with a global personality trait: “I am a positive person”. Also, for North-American participants, a positive attitude was defined as “knowing that I’m capable of doing all the things that I want to do” whereas, for Asian participants, a positive attitude meant “try harder the next time, learn from mistakes, become a better person”.

It is important to note that respondents from Vietnam and the Philippines stated that there was no concept of self-esteem in their language or culture, and the closest construct they had to account for this was the notion of self-respect defined as “holding one’s self accountable to the highest standard”. A distinctive feature of the construct of self-esteem defined this way was the openness and invitation/anticipation of negative feedback from others: someone has high self-esteem if one can accept and integrate negative feedback from others. In addition, for Asian respondents, others and the negative feedback from others appeared to be intrinsic parts of their
self-evaluations and not just as an external reference for upward and downward social comparisons.

**Summary.** In sum, when responding to the RSES items, participants engaged predominantly in self-evaluative and self-regulation processes supported by cognitive, semantic, or affective, self-referenced responding. With respect to self-evaluation processes, participants used agentic-reflective evaluations to complement the self-critical, expectation based self-evaluation. In addition to these response processes, participants also demonstrated engagement in anticipatory-teleological response processes, and social desirable responding for some items. All of these processes seemed to be differentially solicited by different items and it seemed that the negatively worded items tended to engage the self-regulation processes, as well as the emotional and self-referenced responding whereas the positively worded items tended to engage primarily cognitive, semantic processing and positive trait-like self-evaluations. The 4-point Likert-type response scale format was used differently across the RSES items, and some respondents created new response categories (i.e., 2.5 or in the middle category) for some items. In the process of responding to the RSES items, some participants deliberately redefined or changed the meaning of some words and responded to these items according to the new meanings.

**Commentary.** Some of these current findings are somewhat consistent with those of the only study that examined the response patterns on RSES using IRT (Gray-Little et al., 1997) and which indicated that the RSES items do not possess equally effective discrimination power: specifically, items 3 and 7 (positively worded items) appeared to have the highest discrimination capacity whereas items 8 and 9 (negatively worded items) seemed to have performed the worst with respect to this item property. However, the authors have noted that items 8, 9 and 10 might be able to better discriminate at high levels of self-esteem whereas items 3, 5 and 6 seemed to be
more effective in discriminating across all levels of self-esteem. The authors also mentioned the significant impact of qualifiers on the patterns of item responding: “at times” which diluted the item capacity to discriminate because virtually everyone would agree that at times they felt not so good about themselves, and very general qualifiers such as “all in all” or “on the whole” that seemed to have increased the discriminatory capacity of positively worded items.

The results of the present study also suggest that the qualifiers appear to significantly impact how respondents decide to answer the RSES items: “at times” invited an almost immediate moderate agreement with the item content, and tended to override other qualifiers (e.g., “not good at all”), and general qualifiers such as “all in all” and “on the whole” invited a very abstract, impersonal way of responding to items. Although it is impossible to estimate the items discrimination index based on the data of the present study, it appears that the positively and negatively worded items, respectively, seem to be supported by different response processes and pattern of endorsement, and it may be that the differences noted with respect to discrimination power may be at least partially the result of these different processes. This hypothesis needs to be tested in further research studies.

Discussion

This section will discuss the findings of this study using the framework of the model of response processes underlying answering the RSES items, proposed in chapter 2 (*Figure 2.6* and *Figure 2.7*), and will propose some revisions to this model. The explanatory dynamic of the model will also be addressed.

**Epistemological foundation.** The results of the present study can be interpreted in light of the socio-culturally situated, teleological and explanatory-interpretative epistemological stances espoused by the theoretical model proposed in Chapter 2.
**Socio-culturally situated model.** Based on the current findings, the process of responding to the RSES items seemed to be a socio-culturally situated process in the sense that test takers’ answers were significantly shaped by the social and cultural meanings of the self-esteem construct to such an extent that it would be inappropriate to interpret the meaning of the RSES scores without referencing the appropriate socio-cultural meanings and context.

**Teleological model.** The teleological dimensions of the proposed model were reflected in respondents’ frequent references to future goals, wishes and imagined possibilities. Their decisions regarding the final ratings on the RSES appeared to be shaped by these intentional goal oriented and future focused mental processes. Therefore, an accurate interpretation of test scores should refer not only to participants’ remembered experiences and beliefs about their self-esteem but also to their future goals, hopes, wishes and expectations as they shape a positive view of self. The RSES scores appeared to provide an answer to the question ‘what could my self-esteem become’ not only to ‘what is my self-esteem based on my past and present’. In this sense, the RSES scores allow for a dynamic form of assessment of the self-esteem situated in the “zone of proximal development” (Vygotski, 1996, p. 53) or ‘proximal self-esteem’.

**Theoretical model.** The results pertaining to the theoretically expected response processes will be discussed in the next paragraphs using the theoretical framework of the response processes model presented in Figure 2.6.

**Self-referenced responding.** The findings of this study were consistent with the hypothesis regarding the self-referenced responding during answering the RSES items. All participants engaged predominantly in self-referenced responding, either spontaneously or when prompted by the interviewer’s probes. The prevalence of the self-referenced responding was evident in the following instances: the preponderance of the first person reporting, the sustained
access to self-characteristics and autobiographical events, and the spontaneous propensity to compare and contrast self and others (i.e., intersubjectivity). Even the respondents who favoured semantic over episodic processing accessed pertinent self-evaluations and self-characteristics to support their responses to the RSES items. All self-related processes presented in Figure 2.5 were observed when completing the RSES.

These findings suggest that the RSES represents a measure apt to engage self-referenced responding, as expected while completing a self-report measure. A corollary of this is that RSES scores can be interpreted as pertaining to self-characteristics, and self-processes. However, the depth and breadth of the self-referenced responding varied greatly across the items and it seemed to be significantly impacted by some very general or abstract items that predisposed some respondents to a more semantic way of processing and reporting information about self (e.g., mobilizing beliefs and implicit theories about self). Reflecting on this, one of the respondents stated:

Participant: “It’s kind of like a very quick assessment based on my memories – I draw upon them a lot and, uh, I guess I do have a...somewhere in my mind, a compartment of, you know, how I feel about myself, and certain things I’ve decided upon as well, not just...

Interviewer: So you pooled information from there, because you kind of already decided?

Participant: Yeah, yeah, I already knew how I felt about myself. So yeah I already decided and just drew on those memories”

This tendency to resort to semantic processing increases, as expected, with the degree of generality and abstraction of the items. Here is one pertinent comment made by one participant:

Participant: “I feel solid in those items that are asking, you know, sort of overall things. Then I went straight to what I already knew and already decided about myself. The questions
that are more vague, I would maybe answer that differently at a different time depending on what will be on my mind at that time. But these general, overall things will never change: I already decided these, you know, overall things about myself.

Interviewer: Could you tell me what are these overall things about yourself that you feel that will not change?

Participant: I am a person of worth, for example. And I can do things better than others. I am a good, competent worker. I am a kind person. I help people. I like myself. This kind of things. I mean, I know who I am, right?"

Although this tendency towards semantic responding may be due to the very nature of the RSES (i.e., traditional self-report measure targeting beliefs about self), a word of caution is in place when interpreting the RSES scores as state-like, situational or dynamic results.

**Self-evaluation processes.** The results of this study are consistent with the hypothesis that the self-esteem construct depicted in the RSES is essentially a self-evaluative construct. While responding to the RSES items, all participants engaged in self-evaluative processes and strategies in order to come up with their scores on the RSES items (*Table 2.3*), such as: (a) accessing objective self-evaluative information (e.g., test results or achievements), (b) temporal comparisons, (c) downward and upward social comparisons, and (d) searching for or reporting evaluative feedback from others. Overall, a consistent engagement in any or all of the four fundamental self-evaluative processes (i.e., self-assessment, self-enhancement, self-verification or self-improvement) was discernible in participants’ responses. Therefore, these results support the interpretation of the RSES scores as self-evaluative judgements.

**Self-regulation processes.** When responding to the RSES items, the respondents mobilized a multitude of self-regulation processes such as: developing or creating self-esteem (i.e., searching for approval, engaging in value-consistent behaviours), conceptualizing self-
esteem (e.g., developing personal theories based on autobiographical data; organizing the facts in a way congruent with positive self-representations), maintaining and managing self-esteem (e.g., reducing dissonance, refusing to engage in behaviours that contradict the already established self-esteem or personal values, rationalizations, normalization, approval seeking, or success seeking), avoiding further loss of self-esteem (e.g., avoiding challenging situations, over-compensating, complying), enhancing self-esteem (e.g., the pep talk, self-serving bias, emphasis on future possibilities of self-improvement), and restoring self-esteem (e.g., re-engaging in socially sanctioned, value-consistent, and worthy behaviors, self-serving judgments, meaning making in front of threats to self-esteem). The present findings support the claim that the self-regulation processes represent one of the underlying mechanisms of the RSES scores.

‘Self-esteem moments’. The results of this study indicate that, when responding to the RSES items, participants either retrieved an autobiographical event of a ‘self-esteem moment’ (e.g., accomplishments, failure, or helplessness) and used that moment to make decisions about their self-esteem, or they accessed already established declarative self-knowledge relevant to the evaluation of their self-esteem. A spontaneous elicitation of the following autobiographical events was noticed: (a) situation of past failure, (b) situation of past threat to self and self-esteem, (c) anticipated situation of failure or rejection, and (d) a very significant, valuable situation in the present or future.

Some participants provided mainly ready-made responses to the RSES items in the forms of previously formulated beliefs (i.e., semantic responding or declarative self-knowledge) as a reason for their item ratings, especially when responding to the positively worded items. However, there was no clear indication that this predominantly semantic responding represented fake responding or intentionally superficial engagement with the items but rather this tendency
may be suggestive of the difficulties that some general and abstract items have posed to test takers in terms of accessing more personal, concrete content while responding to these items. In this sense, several respondents commented on the difficulty related to making “overall” or “on the whole” judgments about self, and indicated that they responded by accessing what they already decided regarding their self-esteem (i.e., self-declarative knowledge).

**Cognitive-affective self-related processes.** During responding to the RSES items, participants engaged in self-related cognitive processes such as: autobiographical retrieval or episodic memory processes, declarative retrieval pertaining to self-knowledge, counterfactual thinking, and self-attribution processes. In addition, respondents reported experiencing or remembering self-conscious emotions such as: embarrassment, shame and pride. These response processes are consistent with the theoretically expected response processes underlying responding to self-report self-esteem items, and therefore the present results support this component of the theoretical model proposed in chapter 2.

**Response Content.** In the current sample, North-American respondents’ answers accurately and consistently reflected the main dimensions of the self-esteem construct (e.g., competence, self-liking, and self-worthiness). The content of the Asian participants’ responses was not consistent with Rosenberg’s definition and theory of self-esteem. That is, the construct itself was defined differently by these respondents. This chapter will address these differences in a later section about the cultural shaping of the response processes underlying responding to the RSES items.

All North-American respondents referred, in their answers, to most of the main aspects of the self-esteem construct as defined by the theory of self-esteem that influenced the development of the RSES. The most references were to competence and a sense of mastery, followed by
references to relational self-worth, and to self-respect. In addition to the main dimensions of the self-esteem construct as depicted in the RSES, several participants talked more about meaning in life and future plans, compassion and mindfulness, well-being and quality of life, and spirituality when answering the RSES items. These responses may raise issues regarding the construct representation of some of the RSES items and the potential for conflating the self-esteem theoretical dimensions with those relevant to other constructs (e.g., meaning in life).

**Non-verbal processes.** When answering the RSES items, respondents displayed the expected narrative density and coherence characteristic of self-referenced responding in terms of flux, flow, length, richness, idiosyncratic details, and personal authorship. They consistently used the first person framework of reporting and demonstrated the emotional markers associated with affective processes (e.g., self-conscious emotions). Several participants mentioned in their verbal reports the role of “gut feelings”, and the implicit body stored self-knowledge (e.g., “I just know it, I feel it in my body”).

**Summary.** When responding to the RSES items, participants demonstrated most of the theoretically expected response processes included in the model proposed in chapter 2. Specifically, participants’ answers reflected response processes consistent with self-referential, self-evaluative and self-regulation response processes. Participants have also engaged in cognitive-affective self-related processing during responding to the RSES items. The ratio among the autobiographical, episodic retrieval, and semantic retrieval components was, at times, in favour of semantic retrieval for some items, which suggests that some of the RSES items tended to evoke general, abstract evaluations of self. This raises caution about the correct interpretation of test scores as indicative of trait self-esteem and general beliefs about self, and not as state self-esteem measure. In this sense, the RSES conforms to the expected self-processes
encountered in the traditional belief based self-reports. Therefore, the inferences based on the RSES scores should be limited to general self-evaluative beliefs. When answering the RSES items, the Western test takers reported response content consistent with the construct theory such as: competence, values, self-worthiness and likability, social role expectations and identifications, as well as needs, motives, and expectations. However, the content of Asian participants’ responses reflected contents and meanings that were not included in the construct theory underlying the RSES.

The situated, contextualized model. Following the contextualized components included in Figure 2.7, this section will discuss: (a) the cultural variations regarding the definition and meaning of the self-esteem construct, (b) the socially desirable responding when answering the RSES items, and (c) the linguistic aspects such as the wording effect on the RSES.

Cultural differences. The findings of this study suggest that the self-esteem construct is shaped by cultural factors to such an extent that the definition and the meaning of the construct differ cross-culturally quite significantly. The present results are consistent with several research studies that found support for the fact that self-esteem is not a cross-culturally invariant construct (Hyland, Boduszek, Dhingra, Shevlin & Egan, 2014; Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997; Marsh, 1996, 2010).

Feeling good, being good at, and being good for. Whereas the North-American participants included in this study understood self-esteem as a sui generis combination of self-confidence and self-worth accompanied by an overall positive feeling towards one’s self (i.e., self-liking), Asian participants conceptualized self-esteem as the degree to which the individual contributed to collective interests, and the extent to which he or she is good for society and knows how to harmonize with others. Although both North-American and Asian participants
engaged in self-evaluation and self-regulation processes while answering the RSES items, suggesting perhaps an invariant underlying mechanism in generating the test scores, the meaning of these scores on the RSES seemed to differ significantly.

An important difference pertaining to the cultural differences in conceptualizing self-esteem was that, whereas North-American participants described self-esteem as feeling good about oneself, Asian participants understood self-esteem as holding oneself up to the highest standard so that one can honour any promises towards others and contribute to society. In other words, whereas for North-American participants, the self-esteem focus seemed to be on feeling good about self, and others are primarily perceived as external sources of self-esteem, for Asian participants, the focus of self-esteem was on others, and others were seen as an intrinsic part of one’s self-evaluation.

Less is more. These different definitions and meanings of the self-esteem construct appear to have a direct impact on interpreting the RSES scores. For example, Asian respondents’ lower scores on the RSES would mean, in fact, higher levels of experienced self-esteem and self-satisfaction because not endorsing highly individualistic and self-aggrandizing statements is perceived as having healthy self-esteem in the Asian culture, and indicates a moderately high level of self-esteem. This observation is consistent with the research findings about the existence of different self-construals for Westerners and Asians (e.g., Asian people demonstrate a collective self-construal in contrast with the highly individualistic self-construal of Westerners), and with the cultural research about the Asian “modesty bias” encountered in research studies investigating self-esteem (Kitayama et al., 1997).

Self-respect versus respect for others. In addition to the culturally different meaning of self-esteem, the concept of self-respect that is considered an intrinsic part of self-esteem in the
RSES (i.e., item 8) seems to hold different meanings for the North-American and for the Asian participants. Item 8 of the RSES appears to have been problematic in several research studies (Farh & Cheng, 1997; Sedikides, Gaertner, & Toguchi, 2003), and the findings from the current study suggest a possible explanation for that: whereas North-American respondents described self-respect in terms of self-care and self-compassion, the Asian participants defined self-respect as respect for others by not wasting or not taking advantage of their resources. This observation is very important when interpreting the scores obtained for item 8 of the RSES (“I wish to have more respect for myself”) for North-American versus Asian test takers.

A person of worth. Another point of reflection that came up when answering the RSES items pertained to what participants considered as essential characteristics of a person of worth in a specific culture and during a specific historical moment; in other words, what provided the respondents with the basis for deciding on their sense of self-worth? Based on the responses from this research sample, it seems that self-worth stems primarily from personal achievements and material possessions in the North-American culture (e.g., even family and friends tend to be seen as personal accomplishments and extensions of self), whereas for the Asian participants, self-worth appeared to stem primarily from harmonizing with others, and being of value for others and for society. These different ways of defining a person’s worth may lead to different scores on the RSES and to different interpretations of these test scores.

Cultural bias or accurate responding? Farh and Cheng (1997) suggested that culturally Chinese people tended to exhibit a modesty bias in self-evaluation, especially on positively worded items. Therefore, researchers inferred that a wording effect for positively worded items would be evident for culturally Chinese people. However, this may not be a correct assumption: in fact, Asian participants endorsing lower scores on the RSES may be a sign of accurate
responding, not of bias, because of the fundamentally different meaning of the construct. In fact, Asian people who choose moderate ratings on the RSES have a high level of self-esteem as defined in their culture.

A self-critical focus, rather than a positive self-evaluation, may be characteristic of individuals from collectivistic cultures where individuals are primarily concerned with how to fit in, belong, and become part of relevant social relationships (Kitayama, et al., 1997; Sedikides, Gaertner, & Toguchi, 2003). For example, “one’s feelings of worth in Japan may depend less on generally positive self-evaluations and more on self-criticism as a functional means of achieving social harmony” (Heine, Lehman, Markus, & Kitayama, 1999, p. 768). In this sense, empirical studies have documented that mean scores on global self-esteem scales in Japan are from one to two standard deviation units lower than in English-speaking and prototypically individualistic countries such as Canada and Australia (Feather & McKee, 1993). Even more compelling, overall positive self-evaluations tended to increase among Japanese people as their exposure to North American culture increased (Heine et al., 1999). Measures of self-esteem have yielded comparable differences between the individualistic United States and collectivist Hong Kong (Kwan, Bond, & Singelis, 1997) and between the individualistic United Kingdom and collectivist Spain (Tafarodi & Walters, 1999).

These empirical findings suggest that it is possible that not only the meaning of the self-esteem construct but also the specific manner in which the evaluation and self-regulation processes are engaged differ among different cultural groups. The results of the present study suggest that whereas North-American respondents tended to use predominantly self-regulation strategies meant to maintain and restore their self-esteem by buffering the negative feedback and selectively accessing positive memories in order to maintain a positive view of self, Asian
participants preferred to engage in efforts of self-improvement and used motivational self-talk to welcome and respond to critical feedback.

**Conclusion.** These observations raise important concerns regarding the cross-cultural use of the RSES: if used with Asian participants, the meaning of the scores should take into account the cultural meanings of the self-esteem construct. Asian participants may tend to have lower scores on the RSES but these scores are not necessarily indicative of low self-esteem or response bias. Rather, they may be indicative of moderately high levels of self-esteem as defined within the Asian culture: being self-effacing, modest, and conservative in endorsing some items is perceived as a sign of mental health, adaptability and healthy self-esteem in this culture.

These findings suggest the possibility of designing and testing a culturally situated model of self-esteem that could accommodate these different meanings of the construct. Quantitatively, such a model could be a mixture item response theory (IRT) model (Rost, 1991; von Davier & Rost, 1995) that recognizes the qualitative culturally based differences in the content and meaning of self-esteem construct among different cultural groups. Another hypothesis to test in the context of such a model would be whether low and high self-reported self-esteem is also associated with different self-evaluation and self-regulation strategies across cultural groups (e.g., self-improvement versus self-protection).

**Linguistic impact or the “wording effect”**. The hypothesis of a “wording effect” on the RSES came from the observation that the negatively worded items (i.e., items 2, 5, 6, 8 and 9) tended to cluster together and to load on a different factor than the rest of the items (Marsh, 1996; Spector, Van Katwyk, Brannick, & Chen, 1997). Some researchers dismissed this issue and concluded that a tendency for positive and negative items to group into two factors does not necessarily reflect a substantively important distinction between positive and negative self-
esteem (Greenberger, Chen, Dmitrieva, & Farruggia, 2003), and that the differential behavior of positively and negatively worded items may mirror a substantively irrelevant method effect (Marsh). Greenberger et al. (2003) concluded that “one is hard-pressed to argue that the positively worded and the negatively worded versions are tapping into different dimensions of self-esteem” (p. 1252).

However, the findings of this study suggest that, in fact, the differential behavior of the positively and negatively worded items may represent a substantively relevant issue as it does not mirror a mere method effect but different response processes underlying responding to the negatively worded items. Specifically, these findings suggest an alternative explanation for why and how these negatively worded items tend to load on a separate factor: these negatively worded items seem to mobilize slightly different response processes than the positively worded items. Specifically, the negatively worded items were effective at setting in motion a full process of self-evaluation by triggering participants’ ‘self-esteem moments’ such as experiences of incompetence, failure, moral failure, or powerlessness. These autobiographical recollections were accompanied by self-conscious affect that was also differentially shaping respondents’ decisions about the scores via affective processing. Thus, it may not be the negativity of the wording but the content of the personal experiences and the associated processes triggered via item language that led to different loadings for these items. I would argue that these five negatively worded items are actually the most effective items of the RSES in terms of the empirical mapping of the theoretically expected response processes for a self-esteem measure because they appeared to have set in motion the underlying mechanism of self-evaluation and self-regulation by triggering primarily episodic, autobiographical processing, intense self-
conscious emotionality, and motivational processes to regulate self-esteem such as restoring, compensating, and normalizing.

The positively worded items tended to be abstract and general and they were less effective at engaging the respondents personally. Therefore, participants’ responses to these items showed the preponderance of semantic processing, the overuse of declarative self-knowledge in self-assessment and self-evaluation, sporadic self-regulation processes, and overreliance on accessing ready-made feedback from others in coming up with scores for the RSES items. The full theoretically expected response processes model was harder to observe when participants responded to these items, predominantly in the North-American respondents.

These results are consistent with recent research studies investigating the factor structure of the RSES (Hyland, Boduszek, Dhingra, Shevlin & Egan, 2014; Marsh, 2010). These studies have found strong support for a bifactor model: “when examined longitudinally, the bifactor model was found to be the most viable solution, and clear evidence was found of temporal stability within the method factors suggesting that the nature of the wording of the RSES items (positive and negative) represent stable response styles” (Hyland et. al., p. 189). The findings of the present study suggest a similar hypothesis: there may be different stable response processes underlying the negative and positive worded RSES and not just an inconsequential method effect.

Given these results, in the future it would be worth testing a situated response processes model of item wording on the RSES. Specifically, a Person x Situation response processes model can be formulated as follows: (a) in situations of threat to self-esteem triggered by test items, autobiographical and emotional based processing as well as self-protective strategies are used, whereas (b) when answering neutral or positive items, declarative self-knowledge and self-
enhancing or self-verification processes are dominant. The hypothesis is that the accessibility to various sources of self-esteem (e.g., autobiographical, emotional events versus abstract beliefs about self) as well as affective versus cognitive processing would be different across these two situations.

**Socially desirable responding.** At least some of the RSES items appeared to be vulnerable to situational socially desirable responding in the sense of participants attempting to intentionally minimize the reporting of some negative characteristics or to intentionally exaggerate the importance of some positive events or personal characteristics. The positively worded items seemed to be more vulnerable to impression management tendencies for the Western respondents: participants reported examples of good personal qualities, achievements as well as positive attitude towards self. The Asian participants tended to engage in more self-effacing responding when answering positively worded items, and are open to more self-disclosure and engagement while answering the negatively worded items. The situational impression management might have been due at least in part to the interactive administration of the RSES in this study given that the interactive administration also represents a highly activating social situation that may trigger self-esteem moments.

**Revised response processes model.** There were several important response processes missing from the proposed model (Table 3.1), and this section will propose a way of integrating these processes in the revised model of the response processes underlying answering the RSES items.

The scores on the RSES can be seen primarily as the product of two main interconnected underlying mechanisms or response processes: (a) self-evaluation and (b) self-regulation, with respect to several dimensions of the self-esteem construct: a) self-competence (“good at”), b)
self-worthiness (“being good”), c) self-liking (“feeling good” about self), and d) self-
transcendence (“good for”). In the process of deciding for their RSES ratings, respondents access 
pertinent self-esteem related contents such as: self-knowledge, personal experiences, feedback 
from others or imagined future possibilities.

**Response processes.** The primary response processes underlying answering the RSES 
items are the self-evaluation and self-regulation processes. These primary processes were 
supported by several auxiliary response processes such as: emotional, motivational and cognitive 
response processes.

**Self-evaluation processes.** As discussed in chapter 2 (*Figure 2.5*), self-evaluation 
processes consist of several interconnected sub-processes: a) self-assessment (or self-appraisal in 
the language of cognitive psychology) which represents a predominantly cognitive process that 
provides accurate self-knowledge, b) self-attribution processes which are cognitive processes 
that allow one to explain the cause of an event (i.e., primarily self as cause or situation as cause), 
c) self-enhancement processes which ensure the maintenance of an overall favourable self-
esteem, d) self-verification which confers a sense of self-coherence and continuity with respect 
to self-evaluation, and e) self-improvement which pushes self into the future and sees 
possibilities for bettering self and increasing self-esteem.

Self-evaluation is not only a self-critical evaluation with respect to standards and 
expectations set by self or others, but also reflective-agentic evaluation in light of values, 
purposes, and self-transcendence (e.g., “good for”). Whereas the theoretically expected response 
processes model has emphasized the self-critical evaluative processes, the revised model will 
include the reflective-agentic self-evaluative processes as complementary processes.
Self-regulation processes. These processes are set in motion during or after the self-evaluation processes, and are fundamentally concerned with action or behaviours (i.e., what needs to be done to restore, maintain, or enhance self-esteem). For example, once a person decided that he or she made a mistake (self-assessment and self-attribution) that led to activating the need to maintain self-esteem (self-enhancement) to be consistent with his or her view of self (self-verification) and with the potential of getting better (self-improvement), then one may engage in restorative processes (apologizing for mistake), protective processes (defending one’s position or self-esteem), or impression management processes (putting forward a socially desirable response). Self-regulation processes may occur at any point of the self-evaluation processes and inform/shape further self-evaluation processes. Examples of self-regulation processes that were not initially anticipated but should be included in a revised model (Figure 3.1) are: recalibrating the (response) scale of a measure, reshaping the meaning of the construct or words used to describe it, impression management, use of self-serving bias, positive self-talk, and counterfactual thinking.

Auxiliary response processes. The primary mechanisms underlying responding to the RSES items (i.e., self-evaluation and self-regulation) engage auxiliary affective processes (e.g. self-conscious emotionality), motivational-teleological processes (e.g., intentionality), and cognitive processes (e.g. autobiographical memory and declarative self-knowledge). These auxiliary processes are actively involved in making decisions about test scores when answering the RSES.

Sources and contents. The primary and auxiliary response processes recruit their contents by accessing: (a) self-contents (beliefs and theories about self-characteristics, values and expectations) via declarative self-knowledge, (b) self-experiences via autobiographical
memories, (c) feedback from others, (d) objective information (facts, behaviours, achievement), (e) social meanings and contextual cues, and (f) future potential (anticipated possibilities of self-improvement). These contents can also be seen as sources of self-esteem not in a causal sense (i.e., they do not produce self-esteem) but in the sense of providing the contents for the self-evaluative and self-regulation processes. What seems to be “producing” the self-esteem ratings is the interplay of self-evaluation and self-regulation within a social context by recruiting any of these contents depending on their salience and importance for the individual at that moment.

**Dimensions of self-esteem.** The results of the data analyzed in this study support the existence of several components of the self-esteem construct: a) self-competence (“good at”), b) self-worthiness (“being intrinsically good”), c) self-liking (“feeling good” about self), and d) self-transcendence (“good for”). For the Asian respondents “being good with” in the sense of being able to harmonize with others may represent a significant dimensions of self-esteem stemming from the collectivistic socio-cultural stances. This suggests that conceptually self-esteem may not be a unidimensional construct and that respondents refer to several distinct dimensions of self-esteem when answering self-report self-esteem items such as the RSES. Whereas previous research studies distinguished between competence and self-liking as separate aspects of self-esteem, the present study suggests that intrinsic self-worthiness and self-transcendence may also be important facets of the self-esteem construct. Future research studies will need to evaluate this hypothesis.

**Explaining the observed test score variation.** Based on the current findings, I hypothesize that low scores on the RSES may be the result of a combination of the following processes connected with the self-evaluative and self-regulation capacities: (a) ineffective regulation processes meant to restore or maintain self-esteem, (b) intense self-conscious
emotionality overwhelms self-evaluation (e.g., feeling based decisions), (c) over-active self-improvement micro-processes during self-evaluation lead to self-dissatisfaction, (d) overwhelming, unprocessed, unintegrated, and highly invested autobiographical events that render self-regulation ineffective and distort self-evaluation processes, (e) the exclusive presence of critical self-evaluation via expectations and values at the expense of reflective agentic self-evaluation (contribution), or (f) the lack, paucity, or inaccessibility of solid sources of self-validation. It is important to note that moderately low scores for Asian participants are indicative of moderately high self-esteem because of how self-esteem is culturally defined. Therefore, although using the same self-evaluation and self-regulation mechanisms, Asian participants may recruit different contents to support their self-evaluations.

Moderately high scores on the RSES may be due to: (a) strong self-regulation processes able to offset the perceived threat to self during triggering “self-esteem moments”, (b) the predominance of the reflective self-evaluation, (c) the variety, richness, and heightened accessibility of self-esteem sources, and (d) the predominance of situationally negative self-attribution and positive stable self-attributions. Very high scores on the RSES may be due to self-evaluative processes that rely primarily on ready-made beliefs shaped by normative social messages (e.g., one has to be positive, one should have high self-esteem), and to self-regulation processes related to impression management and socially desirable responding. This model is shaped by sociocultural mechanisms

The revised theoretical model of response processes underlying responding to RSES items is presented in Figure 3.1. It is critically important to evaluate to what extent this model can be applied cross-culturally. Although it is possible that the core response processes (i.e., self-evaluation and self-regulation) may be culturally invariant, how exactly are they implemented
and what contents they recruit may be significantly different across cultures. Therefore, a contextualized model of responding to the RSES items would be crucial in order to provide an accurate interpretations of the test scores.
Figure 3.1. Revised Model or Response Processes Underlying Answering to the RSES Items

- **TRIGGERS EVOKED BY ITEMS**
  - Failure
  - Exclusion/Rejection
  - Success
  - Challenge

- **SELF-EVALUATION PROCESSES**
  - Self-improvement
  - Self-verification
  - Self-enhancement
  - Self-assessment

- **SELF-REGULATION PROCESSES**
  - Protect
  - Compensate
  - Restore
  - Enhance

- **POSITIVE FEEDBACK**
- **COUNTERFACTUAL THINKING**
- **COMMITMENT TO SELF-BETTERING**

- **OTHERS**
  - Feedback
  - Social roles
- **SELF-CONTENTS**
  - Beliefs
  - Expectations
- **EXPERIENCES**
  - Past events
- **HOPES**
  - Possibilities
Contributions of the Study

This study makes several contributions to the field of response processes research, and test validity. Overall, findings of this study will contribute to a better understanding of respondents’ mental processes involved in answering self-evaluative self-report measures.

Contribution to the validity evidence of the RSES scores. The current study makes a critical contribution to the substantive validity of the RSES scores. To my knowledge, this study represents the first investigation of the substantive validity of the RSES using a response processes inquiry. Given the widespread use of this scale, in research and in clinical practice, these results are very important given that they help clarify some important aspects related to how to interpret the RSES scores. This contribution can have significant implications for the interpretation and use of the RSES in future research and in clinical practice.

The RSES seems to be a traditional self-report measure of trait global self-esteem that assesses participants’ stable evaluative stances towards self that are mainly grounded in general and already established (ready-made) beliefs about self. Therefore, in spite of several claims that RSES measures state self-esteem, this measure may not be suitable for this purpose. RSES items tend to be acontextual, non-specific, abstract and general, and therefore, are seldom conducive for a sustained personal or affective engagement of test takers. Qualifiers such as: “overall”, “on the whole”, and “all in all” appeared to have pulled respondents into generalization processes and towards a very abstract level of processing and formulating/accessing implicit theories about self and others disconnected from personal experiences. Therefore, although the RSES may be very helpful in providing information about respondents’ general and decontextualized beliefs about their competence and self-worth, it may be largely unsuitable for tracking progress and formulating interventions related to self-esteem.
In spite of being a self-report measure, many times the semantic and abstract processing and responding took precedence over episodic, affective responding. Some rather emotionally laden and highly evocative wording (“no good at all”, “failure”, “useless”) and some radical item wording (e.g., “All in all I am a failure”) seemed to be effective in triggering “self-esteem moments” and the associated response processes, and this type of wording unexpectedly compensated for the highly abstract and acontextual nature of the scale by eliciting a multitude of emotional and motivational processes meant to protect or enhance the threatened self-esteem. This way, the RSES responding appeared to be overall consistent with the theoretical response process model proposed in this dissertation, which emphasized a movement through some well-delineated phases (e.g., self-esteem moments, self-evaluative and self-attribution processes, and self-regulation processes). The content of responses has appropriately matched the main theoretical dimensions of the self-esteem construct defined as a mixture of competence and self-worth. However, it is critical to note that this seems to be accurate only for the Western respondents, and not for the Asian respondents.

Based on the findings in this sample, the RSES scores appeared to be highly culturally and socially sensitive so extreme caution should be exercised when the RSES is used with different cultural groups or in different social contexts. The definition of the self-esteem construct as well as that of related constructs (e.g., self-respect) varies greatly among cultural groups. Moreover the meaning of the construct appeared to be significantly shaped by the social meaning and values poured into the self-esteem construct (e.g., the self-esteem pursuit). Essentially, self-esteem as measured by the RSES represents a construct shaped by the Western way of being and which is non-translatable and not socially or culturally invariant. Nonetheless,
the mechanisms underlying the processes of evaluating self-esteem may be invariant but this hypothesis should be further explored in future empirical studies.

Social desirability and impression management artefacts are likely especially when the RSES is administered orally during a verbal interaction with the test administrator (e.g., several participants voluntarily admitted that they have intentionally minimized the sharing of the full extent of some embarrassing autobiographical event or that they completely concealed any perceived negative personal characteristics in some circumstances so that they were not judged by the test administrator). These observations raise the issues of how to administer the RSES and how to follow up on the test scores before assigning definite meanings and interpretations to the RSES scores.

The cognitive demands on RSES are quite high due to the very abstract and general content of items. This has been apparent in respondents’ very large range of elicited experiences and processes (“On the whole…on the whole, where do I begin? This is a hard one”). These general items may be too hard to adequately process by respondents who may have intellectual disabilities.

In addition, it seems plausible that some of the RSES wording such as “all in all I am a failure” or “I have nothing to be proud of” or “I feel that I am not good at all” are epitomes of some well-documented cognitive distortions and fallacies encountered in most neurotic-level psychopathologies. Hence, this wording may explain the reported high correlations with measures of psychopathology rather than the item content but this hypothesis needs to be tested in further research studies.

**Methodological contribution.** With respect to response processes research, this study represents a methodological innovation by empirically testing a theoretical model of response
processes in order to provide substantive validity evidence. In this sense, this study may serve as an example for future empirical investigations in this area, and it pushes forward this type of research not only by contributing to the empirical findings in this research domain but also by proposing a novel way to conduct this type of research.

**Limitations of the Study**

Being the first of this kind, this study is highly exploratory in its nature. Therefore, this research project should be considered as laying the foundation for possible future research projects in the sense of formulating some more specific hypotheses and theoretical models about answering self-report measures. These hypotheses and models will need to be rigorously tested in future studies before any inferences could be made and applied across various contexts.

The composition of this sample was quite limited in terms of demographic diversity such as ethnicity or even age and gender. Hence, findings of this study will be likely applicable only to similar samples, and no conclusions should be extended beyond the context of this study without additional work to replicate these findings and expand this type of research using different samples and measures.

The method used to investigate response processes (i.e., TAP combined with CI) may not be the most conducive for exploring the situated and socio-cultural dimensions of a response process model. It would be helpful if future research studies could implement a methodology that could intentionally target the socially situated response processes.

In terms of limitations, one may consider that not providing an interrater reliability estimate for the empirical findings presented in this study represents a shortcoming of this project. However, although interrater reliability could have provided additional useful information, I believe that providing such an estimate at this point is not absolutely necessary for
the following reasons: a) the empirical study presented in this chapter is part of an iterative model building process (e.g., the revisions proposed to the initial model), and, thus, expecting agreement between raters at this point would have been premature, b) the main purpose of this dissertation is to propose a proof of concept with respect to a way of theorizing and model building in the area of response processes research and not to determine the raters’ agreement with respect to empirical findings, and c) the epistemological commitments underlying the specific model building strategy presented in this dissertation have intentionally departed from the paradigmatic tenets that require corroboration of evidence for validation of knowledge (Popper, 1934), and, instead, adopted an intersubjective, process-oriented, and constructivist epistemological approach that values subjective and intersubjective meanings, and sociocultural constructive processes as crucial ways of knowledge production (Guba & Lincoln, 2005).
Chapter 4: Observations on Response Processes Research and Its Future Theoretical and Methodological Directions: An Essay

Introduction

The current state in response processes research (i.e., the scarcity of empirical studies on response processes, the methodological paucity with respect to examining these processes, the lack of theoretical models, and the exclusive cognitive focus) is an invitation to critically reflect upon what factors might have led to this situation, and what are the implications of these circumstances on contemporary testing and validation practices, with an eye towards discussing future directions in this research domain. On the heels of this critical reflection, I will then propose that elaborating and advancing a systematic program of research on response processes as well as exploiting the full potential of this type of research may require challenging and expanding some of the currently dominant paradigmatic, epistemological, and practical frameworks. Specifically, I will: (a) explore alternative epistemological and methodological horizons for response processes research; (b) propose the expansion of the scope of this type of research, and (c) discuss potential new roles for response processes research.

A Critical Evaluation of Response Processes Research

Response processes research has not yet flourished and approached its full potential because of a sui-generis combination of factors such as: (a) restrictive overarching epistemologies, (b) limiting theoretical and practical frameworks underlying psychometrics and validity theory, and (c) a certain general climate in social sciences research. Together, these factors have led to a ‘failure to thrive’ in response processes research. This situation has significant consequences for the status of response processes in the day to day reality of validation work, for validity itself, and for developing and validating measures.
Restrictive overarching epistemologies. Applied sciences, such as psychology or psychometrics, are deeply rooted in the epistemological soil that has nurtured them. As Toomela (2009) noted, modern psychology is shaped by a strange mix of two very difficult to integrate epistemologies: (a) the dominant Cartesian-Humean cause-effect epistemology that emphasizes a strict, linear model of efficient causality (experimental research), and/or association (correlational research) accompanied by intensive quantification, and (b) the Aristotelian structural-systemic epistemology that focuses on understanding the structures that underlie behaviours, and on final (teleological, dynamic) causation.

Response processes, defined as the underlying mechanisms that generate the observed test score variation (Embretson, 2010; Messick, 1995), may be considered underlying structures and processes that, through their very nature, would be more compatible research-wise with the structural-systemic (Aristotelian) epistemology than with the dominant Cartesian-Humean epistemology that is intrinsically focused on outcomes and inter-individual differences. Given the supremacy of Cartesian-Humean epistemology over the last century, it is not at all surprising that validity evidence based on response processes has been massively overlooked or replaced by computing associations/correlations among the outcomes of these mechanisms, and that test validity itself has been largely treated as a correlational endeavour based on contiguity and association. What is needed, in my opinion, is to complement this dominant approach with a structural-systemic investigation of response processes underlying test takers’ responses that would allow for alternative causal or non-causal explanations of test score variation.

Restrictive local theoretical frameworks. The ‘breeding soil’ of response processes research has been situated within the epistemological headquarters of psychometrics and test validity. Traditionally, psychometric models and validity theories have been significantly
influenced by several preferred epistemologies, sometimes conflicting with each other, but nonetheless similarly influential: (a) realism (e.g., theoretical entities or constructs, although unobservable, are real), (b) operationalism (e.g., theoretical entities are defined by how they are measured), (c) falsificationism and empiricism (e.g., data fit model approach), and (d) positivism and post-positivism (e.g., true or valid knowledge is accessible via rigorous observation combined with the analytical apparatus of logic and mathematics). Most of these epistemological lenses still exert a considerable influence over today’s psychometrics dealings, and over mainstream ways of conceptualizing validity in testing (e.g., realism- Borsboom, et. al, 2004).

These epistemologies tend to emphasize essentialist, objective, empirical, static, socially disconnected, outcome oriented, deductive, confirmatory, and decontextualized ways of knowing and conducting research. They tend to privilege excessively rarefied and abstract models of reality that, although they may fit well with laboratory experiments and psychometrics models, seem unsuited for understanding and modeling the complexity of living systems processes such as those encountered in the social sciences, and during testing as a social situation.

Rooted in this type of epistemological soil, response processes research has closely mirrored these trends in the sense that it has promoted an exclusively decontextualized, mechanistic, confirmatory, atheoretical, essentialist, abstract, and socially disconnected stance in the empirical investigations conducted in this area. This has led to several important shortcomings in this research domain: (a) significant limitations in scope (i.e., the exclusive investigation of the expected, intra-individual, and cognitive response processes at the expense of emergent, interpersonal, situated, and non-cognitive response processes), (b) disconnection from the essential contextual and social influences that shape any testing situation, (c) dismissal or
minimization of *idiographic* and *qualitative* subjective experiences, and of their associated response processes (e.g., intentionality or agency), and (d) lack of fruitful *theoretical grounding*.

**Inclement climate.** In addition to a relatively inauspicious theoretical soil, some trends in the general climate of social sciences research, in general, and in psychometrics, in particular, have also posed some limitations with respect to response processes research:

First, the “cognitive revolution” (Pinker, 2002) has led to an over-focus on investigating cognitive response processes at the expense of other relevant processes. Although such work is undoubtedly very important, it does not represent the entire domain of response processes. In line with this cognitive focus, the only methods used to examine response processes - cognitive interviewing and the TAP - have targeted exclusively cognitive contents and strategies.

Second, largely atheoretical research practices inspired and supported by dustbowl empiricism and other incentives (e.g., the need to publish quickly) have led to a severe under-development of theoretical models in the social sciences in general, and in the response processes domain in particular. This situation has created the premise for an unprecedented theoretical laziness and starvation in social sciences in general, and for a persistent lack of convincing theoretical models in the area of response processes.

Third, the notable scarcity of research and research frameworks focused on context, processes, and mechanisms has led to ignoring contextual social response processes and to over-focusing on investigating the content of observed responses rather than the mechanisms underlying these responses. Consistent with the focus on outcome research, this tendency has stalled the development of process based models like the ones required in response processes research.
Fourth, in academic research in particular, over-specialization and the pressure to publish finite and nicely packaged empirical findings has encouraged a piecemeal approach to research, devoid of any robust integrative frameworks and cross-disciplinary dialogue. Psychometric work has not always integrated systematically substantive knowledge from various scientific domains, which has led to some excessively thin and, sometimes mechanistic, measurement models. Perhaps an exception would be the construct or model based measurement advocated by Embretson (2010) but, overall, substantive knowledge from various research domains has not been systematically mobilized for the purpose of developing and testing theoretical models.

And finally, the exponential development in statistical technology and the overreliance on statistical models at the expense of field research have encouraged an exclusively nomothetic approach and group aggregate analysis as well as the propagation of very restrictive definitions of constructs that would fit well with the conservative assumptions of most statistical models. It is noteworthy that, recently, some statistical models have become friendlier towards modeling contextual influences, and have also relaxed some of the usual assumptions (Zumbo et al., 2015).

**Conclusion.** The review of the current theoretical and contextual influences on response processes research suggests that what might have appeared initially as a surprising situation (i.e., the scarcity of research studies investigating response processes in spite of their relatively high profile within modern validity theory) represents, in fact, an understandable situation and an almost natural consequence of a more general picture. ‘Thin’ and conservative epistemological frameworks mired in a static, abstract, and socially disconnected worldview combined with dustbowl empiricism and theoretical aversion have contributed significantly to response processes research ‘failure to thrive’.
This state of art has a number of important implications for the status of response processes research in the contemporary practice of validity:

(a) Response processes research has been the ‘Cinderella’ of validity practice, at best an oddity and, at worst, a Pandora’s Box of which any mainstream psychometrician would want to steer clear. Despite the many scholarly articles advocating for the essential role of response processes as substantive validity evidence, these processes have played a rather marginal and ancillary role in contemporary testing and validity practices, and have been largely treated as an expendable and rather costly accessory, a sort of a ‘black box’ between test items and test scores, and, occasionally, as an ad hoc strategy for contextualizing data obtained from testing.

(b) Test takers' response processes have been largely ignored during critical stages of test development and testing. Therefore, important information conveyed by response processes does not explicitly and intentionally inform any of the critical stages of test development and validation, or of an assessment or measurement process. This practical reality is essentially inconsistent with the theoretical stances of modern validity theory that have highlighted that the investigation of these response processes is decisive for understanding the meaning of test scores and for proposing valid interpretations of the test ratings (Bornstein, 2011; Borsboom, 2005, 2013; Zumbo, 2009).

(c) The generative, interpretive, and meaning making response processes involved in answering test items, although essential for a correct understanding and interpretation of test findings, and vital for a correct test use, are largely forgotten, dismissed or, at best, briefly discussed at the end of the testing process. In the relentless pursuit of an ever illusory objectivity, test takers’ subjective experiences as response processes were substituted by ready-made data in
the form of test scores interpreted through a priori lens or a definition of abstract, decontextualized constructs.

(d) By over-focusing on finding evidence for an a priori determined meaning of a construct, the mainstream approach has systematically overlooked the person of the test taker (e.g., the test taker’s subjective experiences, personal life context, intentionality, and agency), and the interaction between that person and the test items within a testing situation (e.g., contextualized assessment), as well as any possibility of novel information stemming from the testing situation itself. Usually, any novel information that does not fit with a priori specifications is dismissed as “construct irrelevant variance” or “bias”.

To a similar extent, the answers on a test are seen as direct reflections or products of the construct of interest and not as meaningful constructions of the person who answered them in interaction with the broader testing context, including socio-cultural and idiosyncratic norms, and personal history. Whereas the construct has taken center stage in validity theory, the test taker who is the generator or creator of the test scores is seen as a passive and largely irrelevant medium through which the abstract construct materializes into numerical test scores that can be later analyzed by various statistical models. The long-standing tendency of ignoring the test taker is in stark contrast with the increased recognition of the social function of testing, and with the advent of the consequences of testing (Messick, 1989; Hubley & Zumbo, 2011). Essentially, testing represents a social activity and a social decision making process. The axiological-pragmatic dimension of testing is a necessary complement to the epistemological function of testing, and the indispensable link with consequential validity.

(e) The impact of language and cultural dimensions on the response processes associated with answering test items have been either largely ignored, briefly discussed at the end of the
assessment process, or addressed as item *bias* under a Differential Item Functioning (DIF) quantitative framework. However, what may appear as bias under statistical inferential ‘logic’ may, in fact, be a crucial aspect of a culturally situated response process cycle not to be dismissed but rather further explored. Multicultural frameworks such as the one developed by Dana (2008) are still marginalized within the mainstream assessment and testing curriculum. Heidegger’s (1951) reflection that human beings dwell in language may be translated into validity language as there is no construct outside language. Language fundamentally determines the meaning and interpretations of test items. Nonetheless, the linguistic and semiotic aspects of response processes have been almost exclusively ignored or have only been explored under a strict and profoundly insufficient cognitive framework (i.e., how do test takers understand the question?).

In conclusion, new ways of conducting response processes research are necessary and they will be discussed in depth in the next section of this chapter. These changes will have to address a new paradigmatic rapprochement by finding new epistemological alliances that could support and nurture response processes research, new methodological strategies, an expanded definition and scope of response processes research as well as reimagining the roles that this research may play in psychometrics, validity, and the social sciences at large.

**Alternative Epistemological and Theoretical Grounding**

**New epistemological alliances.** An important step towards ‘enriching’ the present epistemological soil of response processes research is to examine some more nurturing paradigmatic and theoretical alliances or compatibilities that would further anchor and legitimize this type of research in the bigger picture of contemporary social research and practice. In the next paragraphs, I will briefly discuss some of these possible alternatives as they pertain to
response processes research. An in-depth review of each of these perspectives is beyond the scope of this chapter and the reader is encouraged to consult the suggested references for a more thorough perusal of these approaches.

**Towards paradigmatic complementarity.** Although the Cartesian-Humean and Aristotelian epistemologies seem fundamentally incompatible on many levels, they may be, in fact, complementary for building an integrated, holistic research perspective. Perhaps it is time to remind ourselves that the opposite of a great truth is another great truth, and that these two ‘ways of knowing’ are irreducible epistemological stances that can only allow for fruitful tension and not resolution, and for a new synthesis in Hegelian terms, not for a peaceful, disengaged coexistence.

With respect to response processes research, given its focus on processes and underlying structures and mechanisms, the structural systemic epistemology espoused by the Aristotelian epistemology may inspire some important future scientific strides in this domain. Shifting the epistemological ground in this direction would not mean starting ‘from scratch’ or ‘reinventing the wheel’ but rather intentionally drawing, and building, on the research methods and findings that have been around in psychology for a while, such as: Vygotsky’s (1994, 1996) cultural psychology or Piaget’s genetic epistemology (Piaget, 1970, 1971, 1972). These theoretical approaches align well with the aim of investigating processes and dynamic structures using field data, and integrating intensive qualitative information, theory, and quantitative explorations, such as those required by the research of response processes.

In addition, Lamiell’s (1987, 1990) idiothetic psychology provides an example of paradigmatic complementarity (i.e., nomothetic-idiographic) that could support bridging the qualitative and idiographic aspects of response processes research with the nomothetic features
explored during standardized testing, by investigating the dialectical response processes engaged in when answering test items. The idiographic, qualitative aspects of test takers' responses are relevant for test development and test validity and, in this sense, Meyer, Finn, Eyde, Key, Moreland et al. (2001) highlighted the need for elaborating individualized frameworks for conducting assessments, and emphasized the serious clinical insufficiency and unhelpfulness of practicing assessment from a dominant nomothetic stance. Therefore, a paradigmatic complementarity such as that proposed by Lamiell’s idiothethic framework may be fruitful to this end.

Quantitatively, person-centered modeling efforts that may be helpful in this sense could be the Latent Variables Growth Mixture approach or intra-individual variability models (e.g., idiographic filters and personal factor analysis, Molenaar, 2004; Nesselroade et al., 2007). These models would allow the exploration and integration of both idiographic and nomothetic aspects and would offer a more comprehensive understanding of response processes underlying answering test items.

*The pragmatic-transformative paradigm.* Given its preference for few or no a priori harsh ontological and epistemological claims, and its openness to experimentation, novelty, and social transformation (Merten, 2003), this paradigm may promote a fresher, and more exploratory impetus in response processes research as well as a focus on transformative and socially-meaningful practices in psychometrics and validity in general. For instance, contextual pragmatism (Merten) offers a generous blend of experimentation within socially situated contexts that may support well the examination of interpersonal and ecological response processes.
Postmodern empiricism. The more recent, post-positivist views on empiricism such as constructivist empiricism (van Fraassen, 1980), and structural empiricism (Maturana, 1990, 2009; Varela, 1999) are well-anchored empirically, and, at the same time, have adopted a socially situated, contextualized view with respect to knowledge creation and ways of knowing (Bruner, 1990). These characteristics may support furthering the research on situated, interactive, and collaborative-generative response processes that go beyond exploring the fit between observed and theoretically expected response processes. These epistemologies would fit well with the social transformative stances of the pragmatic-transformative paradigm (Merten, 2003).

Situated (ecological) sociocultural frameworks. By emphasizing the intrinsically social and situated mind, intersubjectivity, dialogicality as well as socio-emotional, motivational, cultural, and identity dynamics (Littleton & Miell, 2004; Moran & John-Steiner, 2004), these theoretical frameworks may encourage the exploration of sociocultural and sociopolitical response processes as they impact the validity of test score interpretation. They could also encourage moving beyond the exclusive focus on cognitive response processes by exploring emotional, motivational, self-related processes as well as meaning making processes. For the latter, Personal Construct Theory (PCT; Kelly, 1963) has explicitly focused on studying test takers’ personal constructs, and their underlying processes, and, thus it may be helpful for exploring the collaborative and shared meaning processes in testing.

Explanatory models: generalization and causation. Generalization and causation are considered critical epistemological processes that are addressed, in one way or another, in every epistemology or theoretical explanatory model. Therefore, it is important to explore in what way these activities may be involved in formulating explanatory models in the area of response processes research. In the next paragraphs, I propose that there are complementary variants of
generalization and causation that may be particularly useful for examining and explaining response processes.

**Analytic generalization.** In contrast with extensive generalization, the key to analytic generalization is the use of “theoretical concepts to enable a more general perspective on specific qualitative patterns” (Halkier, 2011, p. 787). For instance, “in case study research the aim is not to consider the case as a sample of a larger population of like-cases but to discover patterns and processes within the case, and to use analytic generalization to extract the lessons learned” (Erickson, 2012, p. 687). In response processes research, this would mean mindfully recruiting substantive theoretical knowledge that would bring some general coherence and meaning to the patterns identified in the test takers’ scores, even in the absence of a full data set.

Analytic generalization was used extensively by Vygotsky (1994, 1996) and Piaget (1971, 1972) in the area of human development where these scholars focused on exploring and establishing structural patterns using limited field data. These systematized patterns later became the foundations of well-known theories in developmental psychology. Response processes research seems to be uniquely suited for using analytic generalization given the small sample sizes typically recruited in this type of research, the primarily qualitative nature of the data, and the quest for identifying and explaining the underlying mechanisms that could ultimately lead to building theoretical models of response behavior in testing.

**Process causation.** The investigation of response processes fits well with a variant of causality called process causation that is primarily geared towards the study of causal processes (Erickson, 2012, p. 688; Maxwell, 2004, 2012). This is different from the most frequent type of causation implemented in science, in the sense that it uses field-based methods to study specifically and intently the actual array of events and actions that lead to specific outcomes in
local settings (Erickson; Maxwell). Scholars in different fields have recognized the process-focused causality under different names such as: process tracing in political science (Bennett, 2010; Bennett & Elman 2006a; Bennett & Elman 2006b; George & Bennett, 2004; Harrits, 2011), explanation building when doing case studies (Yin, 2014), and process analysis (Anderson & Scott, 2012). The process variant of causality is also flexible enough to bridge between ‘incommensurable paradigms’ (Lincoln & Guba, 2005), and to accommodate the exploration of contextual or ecological aspects.

In essence, investigating response processes means illuminating the processual sequences that underlie answering test items. In-depth observations of these sequences would be geared towards spotting the anomalies in the structure of the qualitative or quantitative data (i.e., response patterns) in order to provide insights into the causal processes of interest. Causal process observations are focused on a causal process, and, thus, are contrasted with the data set observations that are the grist for statistical modeling (Brady & Collier, 2004, p. 338). However, where applicable, “process approaches to causality can have a complementary relationship with typical causal processes within the same study, with the process inquiry producing additional (process) information that is absent from what is essentially an input–output analysis between the independent and dependent variables” (Harrits, 2011, p. 153).

**Agentic or final causation.** Lincoln and Guba (2005) pointed to the paradoxical situation wherein “humans, being anticipatory, can produce an effect in anticipation of its cause” (p. 142). The idea of final causation or teleology is not new but rather a few millennia old (i.e., Aristotle). In the previous century, Heidegger (1951) also wrote extensively about “causa finalis” as a distinct type of causation that takes pre-eminence in human dealings due to human’s capacity for intentionality and anticipation. Agentic or final causation would be particularly relevant for
investigating and explaining the agentic and teleological response processes that may shape the test scores during a testing process.

**Theoretical models.** A careful integration of empirical knowledge and theory is crucial for building sound theoretical models given that theory building can go astray if previous substantive knowledge is not properly integrated into the models, whereas the lack of empirical grounding may lead to empty and indefensible theories (Hesse-Biber & Burke Johnson, 2015). In response processes research, this is a critical issue given the extensive lack of theoretical grounding and/or theoretical models or frameworks.

A theoretical model of response processes could be built by triangulating qualitative and quantitative information, and following a sequential interplay of inductive (data driven) and deductive (theorizing) stages. For example, formulating empirically grounded theoretical hypotheses from the observed response patterns in the qualitative data could be the beginning of the model building process. In the next step, relevant substantive knowledge may be mobilized to consolidate or alter these preliminary hypotheses, and new theoretically based hypotheses could be generated and further tested with additional qualitative or quantitative data. Alternatively, substantive knowledge may be the impetus and or the starting point for generating hypotheses that can be further clarified using qualitative or quantitative data. Either way, a purposeful integration of theoretical and empirical rationales is critical for the building, testing and refining of a model. This type of model building represents more than simply checking if the observed response processes fit with the theoretically expected response processes (Standards, 2014). The emphasis is not on assessing the fit with a pre-established theoretical construct assumed to be the standard but rather on progressively building knowledge by a back and forth interplay between theoretical rationales and empirical data.
In addition, a theoretical model of response processes should include an explicit account about how the components of the model interact with each other; in other words, it should clearly explain how the model works. A good model would also be flexible enough to situate itself well across different setting or testing situations. Ideally, a response processes model would be a local and flexible model, not a grand theory. Although there seems to be a preference for developing explanatory models, at this stage, robust empirically grounded descriptive models would be equally valuable for the purpose of advancing the understanding of the relevant response processes that shape the interpretations of test scores. Explanation and understanding are not competitive or mutually exclusive activities but rather necessary complements.

**Methodological Horizons**

Methodology contains more or less explicitly a worldview, an anthropology, ethics and politics:

Methodology—contrary to currently accepted views that consider it a ‘toolbox’ of ready-to-use and consensually ‘certified’ methods—is a process of meaning construction that entails all aspects of scientific inquiry. It is a cycle—or helix— that entails viewing underlying assumptions and the theoretical implications that grow out of them, while in parallel examining the immediately available aspects of the phenomena [...] (Valsiner, 2003, p. 187).

These assumptions have to be intentionally examined and transformed if necessary in order for response processes research to flourish.

**Explanation and description.** Although the quest for explanation and explanatory models are laudable in any scientific endeavor and even the scientific aim per se, in response processes research there has been a rather premature push and exclusive impetus towards
formulating neat and tight explanations, either deterministic or contextual, even before earning a good, solid grasp of the multifaceted reality of response processes associated with answering test items. As already discussed on several occasions in this dissertation, very few empirical investigations on response processes have been conducted, and this domain is far from being thoroughly researched and understood at a robust descriptive level. Moreover, there are no clear or comprehensive methodological frameworks about how to research response processes other than automatically resorting to conducting cognitive interviewing or “think aloud protocol” sessions.

Therefore, eliciting rich descriptions of the response processes involved in testing via complementary qualitative methodologies would be helpful in order to better understand the subject matter, not only abstractly but also as a situated and socially relevant phenomenon. For example, discourse analysis (Derrida, 1976, 1982; Foucault, 1972, 1991), which is focused on eliciting meaning making processes from narratives while taking into account how language is shaping and building understanding and meanings, may be particularly suited for investigating response processes that emerge from test takers’ narratives during qualitative interviews about how test takers relate, and respond, to test items. The Listening Guide (Gilligan, 1982; Gilligan, Spencer, Weinberg & Bertsch, 2005), an analysis that focuses on voice, resonance, and relationship as entry points into the human psyche, may also support understanding the polyvocality (and sometimes contradictory nature) of test takers’ response processes, and contribute to an in-depth understanding in this research domain.

Relaxing the confirmatory impetus. The idea that there is a conceptual yardstick (i.e., construct) against which one compares and measures the empirical data in order to explore the “best fit” represents a statistical reflex that may do more harm than good at this stage in the area
of response processes. Achieving the best fit model works effectively in inferential statistics, but it has the potential to stall progress in the area of response processes research because it forces researchers to operate selectively and conservatively in a field with very few empirical data and almost no models to test. In my opinion, it is far too early to look for confirmations before we even know what are we looking for, and what are we going to find. In this sense, Tukey (1986) stated with respect to the state of art in psychometrics: “Exploration has been rather neglected; confirmation has been rather sanctified. Neither action is justifiable.” (p. 822). This remark is a good reminder to keep the balance, and to temporarily relax the stringent confirmatory and explanatory discourses, at least until we know well enough the object of this type of discourse.

That being said, it is certainly not too early, but rather long overdue, to start generating hypotheses grounded in rich empirical data and supported by theoretical rationales, to build local models, and to experiment with these local models while bringing in substantive knowledge to support the conceptualization process. At this still incipient stage, more effort and attention should be dedicated to the generation of hypotheses and data exploration rather than examining if the empirical data fit with the construct theory. Although it is often forgotten, “hypothesis generation is a crucial stage of research because good experiments test specific and informative hypotheses” (Freedman, 2010, p. 49). Working exploratorily from the data to theory and back would be more helpful than focusing on testing models in a top-down, confirmatory fashion.

“Thinking outside the Q box” (Pearce, 2015). We are past the time when we can afford to contrast and confront quality and quantity (“the Q box”), or to pick and choose whichever fits with our idiosyncratic epistemological landscape. Any responsible and credible scientific endeavour must be fully conversant in both languages as scientific inquiry is essentially bilingual: “Scientific inquiry is a long and tortuous process, with many false starts and blind
alleys. Combining qualitative insights and quantitative analysis—and a healthy dose of skepticism—may provide the most secure results” (Freedman, 2010, p 352). In this sense, Hegel’s (1977) remark that measurement represents the unity between quality and quantity is a timely reminder: quantities are characteristics of a quality, and until quality is elucidated and established, quantitative descriptions would be largely meaninglessness. It is ironic, although certainly not overly surprising, that historically the quantitative imperative has dominated sciences in the last century or more.

For instance, obtaining substantial knowledge about the response processes involved in testing requires a purposeful immersion in the data since “no amount of statistical maneuvering can get very far without a deep understanding of how the data were generated” (Freedman, 2010, p. 23) or of the qualitative structure of the data. Researchers who rely on observational data, as is the case with response processes, need qualitative and quantitative evidence, including case studies. Therefore, integrating complementary qualitative and quantitative frameworks in researching response processes seems to promise the most solid results.

It is also critical to recognize that, from a quantitative standpoint, the recent psychometric developments in the area of latent class mixture modeling (LCA) allow for a more flexible inclusion of context in statistical models and for modeling qualitative heterogeneity and diverse response patterns clustered as latent classes (Sawatzky, Ratner, Kopec & Zumbo, 2012). One way in which response processes research will benefit is to quantitatively model and test hypotheses generated from qualitative data using LCA frameworks.

**Integrating inductive and deductive frameworks.** Working exclusively from a top-down perspective (i.e., a theoretical model applied to data) may be too conservative and would constrict the richness of the data in this relatively new research domain. Working exclusively in
an inductive fashion would lead to a piecemeal type of research, and to no explanatory or
descriptive models. Therefore, a dialogue and mutual shaping between both strategies would be
ideal. The inductive emergent trend will be best suited for understanding contextual, situated
response processes whereas the deductive direction would work best to mobilize existent theories
for making sense of data. After an inductive immersion in the data phase, new theoretical efforts
should be put forward in the form of preliminary hypotheses and, then, a new immersion in data
guided by these hypotheses can follow. This deductive-inductive interplay also fits well with the
need to balance confirmatory-exploratory and explanatory-descriptive tensions. Building situated
response processes models requires working at the intersection of multiple polarities or tensions:
deductive-inductive, confirmatory-exploratory, and explanatory-descriptive. Hence, coding
schemes should be flexible enough to allow for a continuous scaffolding of processes as the
researcher advances through the data and to resist premature “closure” for the sake of
formulating the best fit model. A model should be theoretically informed but not rigidly theory
driven given that it is possible that new theoretical insights emerge from the empirical
investigations on response processes.

**Mixed methods designs.** This type of design lends itself well to integrating apparently
contradictory empirical findings or even epistemological stances, such as those that are likely to
be encountered in response processes research. Mixed methods designs “enable a progressive
reconfiguration of substantive findings and interpretations in a pattern of increasing insight and
sophistication” (Caracelli & Greene, 1997, p. 23), and, therefore, may support the development
of empirically grounded knowledge in connection with theory. There are several types of mixed
method designs that could be helpful for researching response processes (e.g., triangulated,
embedded, exploratory, explanatory, and iterative designs). An exploration of each of these is
beyond the scope of this paper; however, it is noteworthy to remark that potentially all of these
designs can support an ongoing integration of quantitative and qualitative findings and make
room for rich or ‘thick’ methods or methodologies that may enrich the empirical and situated
knowledge on response processes.

**Principled discovery.** Mark and colleagues (Mark, 2003; Mark, Henry, & Julnes, 2000)
have coined the term ‘principled discovery’ to describe a set of methods that can allow for
emergent revisions and elaborations of hypotheses, but with procedures to reduce the likelihood
of being misled by chance (i.e., statistical models). In essence, principled discovery involves two
primary steps. First, the researcher carries out some form of exploratory analyses that may result
in a finding that goes beyond the initial a priori hypothesis. This discovery may point to an
underlying mechanism (i.e., mediator). Therefore, the second general step of principled
discovery requires the researcher to seek some form of independent (or quasi-independent)
confirmation of the discovery, either quantitatively or qualitatively, together with mobilizing
relevant theoretical principles that may be relevant for making sense of that discovery. In
response processes research, this may mean that, after discovering patterns in the qualitative data
of test responses, researchers may employ theoretical principles and further empirical
explorations to formulate an understanding or explanation of that circumstance. This would
involve an intentional revisiting and re-examining of the data, a situation about which Rosenthal
(1994) wrote: “Many of us have been taught that it is technically improper and perhaps even
immoral to analyze and reanalyze our data in many ways (i.e., to snoop around in the data). We
were taught to test the prediction with one particular preplanned test . . . and definitely not look
further at our data. . . . [This] makes for bad science and for bad ethics” (p. 130).
In conclusion, response processes research can draw on several integrative methodological frameworks that could be generous enough to allow for an in-depth qualitative exploration of test takers’ response patterns, and, at the same time, are rigorous and open for dialoguing with theoretical rationales and quantitative analyses to build descriptive and explanatory models in this field. Hopefully, these methodological foundations would also allow for a burgeoning of the empirical studies in this area.

**Alternative methods.** Diversifying the methods of investigating response processes beyond cognitive interviewing and TAP would mean, in essence, to expand the space between items and responses to examine what is going on within that generative space. Rather than focusing on the final ratings or test scores, the focus should be on how these ratings are generated or what happens between items and response. In the next paragraphs, I will suggest some methods that could be fruitful to this end.

**Alternative test items and response formats.** Building tests items that are versatile enough to allow the dynamic co-construction of meaning, and the emergence of response processes during testing would be an important methodological innovation with respect to how to investigate the response processes underlying responding to various items. Specifically, this means developing different types of test items and response formats that could offer test takers the chance to interact with items and to select/modify items in real time. Some examples could be: a) using modified reconstructed response formats for self-report measures (Valsiner, 2005) that would allow test takers to define the meaning of the anchors on a Likert type response scale and, if needed, recalibrate the response scale (e.g., “false” may be defined differently by test takers either as “never happened to me”, “I don’t remember” or “this makes no sense to me”), b) developing and administering constructive items to complement the test administrator’s ready-
made test items (e.g., laddering or repertory grid), and c) including respondents’ constructions of a construct as part of testing (e.g., asking for participants’ understanding of the construct of interest, either during the test development phase or during validation studies).

Below there are some examples of how these strategies could be used when measuring/assessing the self-esteem construct to complement and expand the information provided by traditional self-esteem measures. In the case of the Rosenberg Self-Esteem Scale specifically, one could intentionally elicit (a) the meaning of the items by inviting the test takers to explain what each item means to them, b) test takers’ comments or suggestions about each proposed item of the scale by providing a space next to the items where test takers can write their impressions, thoughts, suggestions or feedback about that item, (c) the meaning behind the item response formats by encouraging test takers to define for themselves what each anchor of the response format scale means, and under what circumstances would they use each of these categories, and (d) the meaning of the targeted construct and items by using interviewing techniques such as the repertory grid to uncover test takers’ personal constructs network that is mobilized when responding to an item (e.g., for the item “I am a person of worth” respondents could be asked the personal constructs that they connect with “worth” and “person of worth”).

**Qualitative comments and feedback.** Inviting and collecting respondents’ commentaries while they answer self-report items about the meanings of the items, questions that are hard to interpret, or meaningless items (e.g. “Please describe your thoughts and feelings about the item/question/statement”, Valsiner, 2005), as well as soliciting immediate feedback about the items (e.g., “How is it for you to attend to this question or to reflect on this statement?”) could be ways to intentionally explore the response processes underlying answering test items. Then these qualitative data would be analyzed in conjunction with the quantitative patterns of test scores.
**Meaning focused methods.** Using real time methods of investigating the ongoing emergence of meaning during the testing process, such as the microgenetic or aktuelgenesis method (i.e., studying the emergence of new meanings when answering test items; Wagoner, 2009), semiotic scaffolding (Hoffmeyer, 2013), or methods of experience sampling in real-life settings (Mehl & Connor, 2012) would allow one to investigate the emergence of meaning and the study of meaning making as response processes.

**Memory schema reconstruction.** Traditional self-reports target aspects of the remembered self (i.e., experiences and beliefs that test takers remember about themselves). Bartlett (1983) developed a method of investigating remembered aspects of one’s life as a reconstructive process. Building on this idea, Wagoner (2009) suggested ways of studying remembering as a contextualized and constructive process. This methods may be fruitful in unpacking the response processes underlying self-reported information about the remembered self.

**Experimental methods.** Using experimental methods such as experimental manipulation of test items in different testing situations would allow one to uncover possible prototypical response patterns in specific situations and to discern how items are interpreted in various contexts.

**Integrating sources.** One could integrate systematic behavioural observations during testing with biometric data (e.g., eye tracking, response time, heart rate, skin galvanic response, and cortisol levels) and fMRI data to identify patterns of responses.

**Defining the field of meaning.** Responses to test items need to be grounded in a field of meaning that is contextual. Therefore, it would be important to either provide a definition of the context/meaning or to allow participants to set their meaning anchors or to define their field of
meaning, and focus on a uniform interpretation of test scores, not on uniform wording and administration given that the same word may be interpret differently by test takers.

**Expanding the Scope of Response Processes Research**

This section will explore the possibility of expanding the investigation of response processes beyond: (a) a cognitive focus, (b) an intra-individual focus, and (c) a theoretically expected focus, in order to examine (a) conative and self-referential, (b) contextualized and situated, and (c) emergent response processes.

**Beyond cognitive response processes.** The emphasis on cognitive response processes has led to a nearly complete overlooking of non-cognitive response processes in research studies. **Cognitive response processes.** The most researched response processes underlying testing tasks are cognitive processes, such as: item comprehension, interpretation, retrieval, and expression of an answer to test items (Tourangeau, 1994; Schwarz, 1999), information processing and problem solving strategies (Embretson, 1983, 2010), and the cognitive operations involved in dealing with cognitive complexity (Embretson & Gorin, 2001). In addition, most of the existing guidelines about how to study response processes (Standards, 2014) have focused on the cognitive processes, and have recommended cognitive interviewing as the method of choice to tap into these processes (Wills, 2005). Although there is no doubt that cognitive processes, such as understanding and interpreting test items are critical, and are likely to be involved to a certain degree in all tests, cognitive processes are not the only type of response processes that underlie the production of test scores, and that are relevant to the validity of inferences made from test scores.

**Conative response processes.** When answering test items, test takers are not only cognitively engaged with these items but also emotionally and motivationally engaged; the
process of generating test answers is strongly connected with evaluative and self-evaluative processes (e.g., importance, valence, and meaningfulness), motivational processes (e.g., motivations around impression management when answering test items), reflexivity and self-referential processes, and meaning making processes, to name but a few.

**Self-referential response processes.** A specific category of response processes that are primarily involved in answering items on self-report measures are the self-referential processes. Given that self-report measures form the majority of tests in some areas of psychology, the lack of research about these response processes is quite stunning. The way in which test takers access self-related information, and link information provided on test items to self is qualitatively different from how test takers process the cognitive aspects of items, and different from the response processes involved in problem solving. In Chapter 2 of this dissertation, I addressed the self-referential processes involved in self-report measures, and in Chapter 3, I discussed the findings of exploring self-referential response processes in an applied context.

**Teleological, agentic, and intentional response processes.** Far from being a mere reflection or a causal product of a construct, the result of any assessment process resides in the space between the item and a response, not solely in the test item, and not solely in the score (Bazerman, 1995, Markus & Borsboom, 2013). This generative space between test item and test score represents the focus of response processes research. Paraphrasing Frankl’s (1982) remark that: “between stimulus and response there is a space and in that space lies the human freedom” (p. 56), we can say that between the test item (stimulus-S) and the test score (response-R) lies the freedom of generative response processes, of human intentionality, agentic reasoning, and decision making processes that cannot be reduced to a simple and direct link between S and R.
Test takers’ subjective experiences, intentionality, and meaning making processes have been largely ignored by the S-R paradigm, and by the “best fit” approach in testing. It is time to reclaim these processes at the center of validity. What has been, for too long, considered to be the black box of testing and validity may well be a treasure chest where we could find the key to a more accurate, contextual, and meaningful interpretation of test scores.

**Meaning making response processes.** Making sense and making meaning are powerful human motivational processes. Human beings are powerfully motivated to bring coherence to their world or to make sense of it, and to find or construct meanings. Testing is an intrinsically meaningful activity; it has a purpose, uses language (specifically, two languages: numbers and words) and language interpretations to make sense of its findings, conveys ample interpretations of test scores and inferences, and is situated at the confluence of multiple value systems. In this sense, each test item can be conceived of as an invitation to “meaning making”, a sign around which a web of meanings is created through a series of iterative processes that attempt to decide which answer makes sense when certain contexts are invoked. Although meaning making processes are critical for testing, they have been the least explored and the least understood (Markus & Borsboom, 2013).

**Meaning and context.** Everyone who has ever given or taken a test probably remembers the most frequent unsolicited comment in response to more general test items: “it depends”. This can be seen as one of the markers of a contextualized meaning making process. Each respondent makes sense of the items by situating these items in a personally relevant context (e.g., by relating items to one’s own beliefs, worldview, emotions, personal experiences, or in comparison to the others one knows). In addition to the semantic understanding of test items, there is a *pragmatic* function of answering test questions that is connected with the context of testing.
Moreover, there is no meaning or meaning making without a context (Leontiev, 1999, 2014), and the human mind is constitutively contextualized: “It seems reasonable to abandon the simplistic version of ecology—minds in context, in situational surroundings, and to pursue a more constitutive study of ecology. Mind can be studied as intrinsically social and contextualized” (Edwards & Potter, 1995a, p. 35; Edwards & Potter, 1995b, p. 36). This means that context and contextualized response processes are central for testing, not external variables, not threats to validity (Standards, 2014) but rather an “ecological niche” (Cole, 1985, p. 74).

**Beyond an intra-individual focus towards intersubjective and interactionist frameworks.** One of the consequences of the severe individualistic bias within contemporary methodology (Farr, 1998) is the absence of a sustained and interdisciplinary discussion regarding appropriate methodologies for studying inter-subjectivity. Even in research on intersubjectivity, the unit of analysis is often the individual (O’Donnell, Tharp & Wilson, 1993). Nonetheless, testing represents an interactive situation in which the test taker interacts not only with the test items but also with a specific testing situation and with the test administrator (or with the computer setting of test administration). Therefore, testing or answering test items can be seen as a type of a person-situation interaction (Mischel & Shoda, 1995).

If we see testing this way, then it is natural to see that test takers’ response processes emerge at, and from, the intersection between the person who is taking the test and the situation defined by the test. For example, several research studies noted that test takers’ scores have changed significantly after experimentally manipulating the valences of the testing situation (Holtgraves, 2004). Specifically, a series of motivational and evaluative processes are set in motion at the intersection between a testing situation and test takers. This has substantial impact on the test scores and, further, on the interpretations of these scores.
In order to examine and understand these situational response processes that arise at the intersection between the test taker and the test situation, we can use Mischel & Shoda (1995) social cognitive-affective theory stating that:

Understanding individual functioning requires identifying the psychological situations that engage a particular person's characteristic/representative personality processes and the distinctive cognitions and affects that are experienced in them. Then, an individual's functioning should become visible in the distinctive or unique ways the person's behaviour is changing across situations, not just in its overall level or mean of functioning (p. 674).

Testing represents precisely such a situation. Thus, it becomes important to identify what kind of psychological responses are evoked by certain test situations, and what are the situational and situated prototypic behaviours/responses engaged by test taker in those specific situations.

Furthermore, we can regard situational response processes as "stable but discriminative patterns of behaviours across situations or as unique bundles of temporally stable prototypic behaviours contextualized in psychological situations" (Mischel & Shoda, 1995, p. 674). In social cognitive theory, individual differences in patterns of behaviour across situations reflect underlying person variables such as an “individual's construals of their experiences, expectations, goals, values and self-regulatory strategies” (p. 675). The individual construals are very much an analogue to Kelley’s (1955) personal constructs theory. These interactive construals are all critical in shaping test scores and in conveying meaning to test ratings, and they fit very well with a contextual, pragmatic explanation framework (van Fraassen, 1980; Zumbo, 2009).

Epistemologically, research on situated or contextual response processes can be conceptualized as an "idiographic analysis of behavioural coherence" across testing situations (Mischel & Shoda, 1995, p. 675). Thus, investigating situated response processes is critical to
understanding intra-individual variability and patterns over time, and it represents the necessary complement to the nomothetic, aggregate based interpretations that most tests yield just by the way they were constructed. This endeavour would also allow one to refocus testing and measurement on the intra-individual pattern of variability in addition to the "broad overall average individual differences at aggregate level" (p. 685). At the same time, a situated, interactive approach could avoid the danger of solipsism that ignores any universal, shared meanings and any possibility of exploring nomothetic relationships/laws.

**Dialogical response processes: Testing as conversation.** In order to expand the domain of interactive response processes in testing, we can conceive of testing as a conversation or dialogue (Markus & Borsboom, 2013; Westerman, 2003). Markus and Borsboom have proposed to contrast the test seen as a detached observation of a construct with the test conceived of as a conversation, with test questions and answers forming a sui generis dialogue. If we further examine this contrast, we begin to see that testing as observation renders test users to a passive stance of observing and recording test ratings, and then interpreting the test scores as reflections of constructs. Testing as conversation involves an interactive attitude and a shared understanding of the items asked and answered. The latter is a conversation that grows organically and is situated within a community of shared meanings that transcend the predetermined meaning of the given construct. How the conversation unfolds cannot be predicted from the onset as it springs beyond the classic data-model fit into yet unexplored territories. Testing as conversation requires a courageous and a rather uncomfortable relinquishing of validity conceived on a fixed, predetermined foundation in favor of an indeterminate and open ended process.

The metaphor of testing as conversation was also suggested by Westerman (2003), and Schwarz (1999), who suggested conversation analysis as a possible method for studying
interactive, dialogical response processes involved in answering items sometimes iteratively, over several regulatory loops. Identifying the main interactive processes may help by incorporating the relevant ones in the test development/measurement model. Giorgi’s (1970) method of analyzing dialogue may also be explored as a way of uncovering the dialogical response processes involved in testing. Specifically, Giorgi developed a phenomenological method for analyzing the structure of conversations, and some of these techniques could be applied to clinical interviews and testing is testing is seen as a conversation or dialogue.

**Beyond inter-subjectivity towards socio-cultural, socio-political, and linguistic frameworks.** The response processes associated with testing occur not only at intra-individual or interpersonal levels, but also, within and in dialogue with, the larger sociocultural context and shared linguistics frameworks that define or shape the meaning of test items and, implicitly, of test scores. In addition to individual, idiosyncratic meanings, there are meaning structures that are located within complex societal relationships. These meanings shape test takers’ answers to test scores as much as they shape test developers’ definition and theory of a construct. For example, the self-esteem construct does not have only an individual content and meaning but it is actually significantly shaped by socio-cultural dynamics that influence test takers and test developers alike (e.g., the pursuit of high self-esteem, virtually everyone sees their difficulties in terms of low self-esteem, society and self-help books regulate what is a high enough level of self-esteem, test takers’ worldviews include self-esteem as a mandatory requirement for a good life or as a sort of ‘social vaccine’). All of these are not just individual propensities, and it would be incorrect to interpret the response processes that underlie the test scores only as construct specific underlying mechanisms.
Trans-individual aspects in testing should be regarded as social, cultural and political ‘negotiations’ of meanings and actions (Gergen, 2009, Shotter, 1993) while coming up with an answer or rating. Socio-cultural lenses applied to testing refer to how inferences and interpretations of test scores are socially negotiated at the intersection of multiple realities. Explicitly formulating these shared social meanings allows for a valid argument for test score interpretations. Test score interpretations are never only about the individual who answered the test but about that person as part of the social fabric. Testing and methodology are political stances, not mere dispassionate, ‘objective’ science.

**Collaborative meanings.** Test developers’ meaning of a construct is not always shared by test takers. Therefore, it is not the examination of the test or of the task itself that is the most important for providing validity evidence but rather *what test takers make of it* and *what raters make of what test takers made of it*. Thus, a main focus in the future response processes research would be to examine the *collaborative* and sometimes shared *meaning making processes* that take place during testing, and how they impact the test scores. Moreover, to paraphrase, meaning is not only in the test taker’s head, and does not depend exclusively on test takers’ interpretations but “it is partly determined by (a) the structures of the world itself, and (b) the linguistic conventions of the social community” (Markus & Borsboom, 2013, p. 507). If we assume that meaning is not entirely in the head, then, when we develop validity arguments, we have to take into consideration the socio-cultural context and the response processes that stem from this context as much as we are willing to examine intraindividual response processes.

**Linguistic frameworks.** Bakhtin (1990) has noted that we ‘rent’ our words from their prior users and uses, extending them to those with whom we currently interact. Thus, the meaning of test scores cannot be separated from the linguistic conventions and interactions in
which it acquires its significance and relevance. Moreover, testing means being fluent in two languages (words and numbers), and able to gracefully translate between them. Potter and Wetherell’s (1987) book *Discourse and Social Psychology: Beyond Attitudes and Behaviour* has marked the so called ‘turn to language’ undertaking in psychology, and it has challenged the knowledge and assumptions of cognitivism and experimentalism that dominated psychology at the time by providing a reconceptualization of language as a means of constructing social reality and achieving social objectives. The overwhelming majority of response processes underlying testing are linguistic in nature, and, thus, it makes sense to investigate these linguistic mechanisms as part of response processes research.

**Moving towards emergent response processes.** So far, the guidelines pertaining to response processes research have focused exclusively examining whether the observed response processes match the theoretically expected response processes (*Standards*, 2014). Whereas this focus is undoubtedly very important, it is equally important to recognize that, in testing, there may be relevant response processes that are not predictable from the beginning solely on the basis of the construct theory/meaning. Nonetheless, these as of yet unpredictable and undiscovered response processes shape the test scores and may be equally relevant to the inferences/interpretations made from test scores.

If testing is seen as observation in light of a construct theory, then we only focus on expected response processes as predicted by the theory of the construct and relevant, auxiliary, ad hoc theories (e.g., theories of self and self-related processes). If we approach testing as conversation, then we become interested not only in the expected response processes but equally interested in the emergent, situated, constructed response processes that are seldom fully predictable from the beginning of testing only from the construct theory. Nonetheless, these
emergent response processes may significantly impact the interpretations made of test scores, and, hence, have to be investigated by the response processes research.

**Conclusion.** When discussing response processes in testing and as source of validity of inferences made from test scores, it is important to distinguish between: theory expected response processes (based on construct theory), theory expected processes based on psychological knowledge/empirical findings, and emergent, interactive and situated response processes that are unpredictable based on construct theory. The latter pose a significant challenge to validity theory focused on construct and matching empirical findings with construct because there is no room for emergent processes in this view. However, if we open and broaden our view on validity beyond fitting data to theoretical constructs, to emphasize the situated, interpretative pole of validity, then we can imagine an indeterminate, fluid validity that incorporates construct theory as a very important component but extends beyond that towards a contextualized, fluid, person-centered validity.

**Reimagining the Roles of Response Processes**

The next paragraphs will examine the main roles or functions that response processes may fulfill in the field of psychometrics and test validity. It has been recognized that response processes represent one of the main sources of validity evidence (*Standards*, 2014). In addition to this, response processes may play other roles that can contribute to strengthening validity arguments. *Table 4.1* presents the main roles that response processes may play in validity and testing.
Table 4.1

Reimagining the Roles of Response Processes

<table>
<thead>
<tr>
<th>Role</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmatory</td>
<td>Examining the fit with the theoretically expected response processes</td>
</tr>
<tr>
<td>Exploratory</td>
<td>Exploring and generating new information relevant to testing process and validity</td>
</tr>
<tr>
<td>Psychometric parameters</td>
<td>Mathematically formalizing some response processes and including these in psychometric models</td>
</tr>
<tr>
<td>Feedback loops</td>
<td>Actively informing and shaping the development, piloting, and refining of test items</td>
</tr>
<tr>
<td>Research variables</td>
<td>Experimentally investigating and manipulating response processes across different testing situations, and over time</td>
</tr>
<tr>
<td>Theory building blocks</td>
<td>Contributing to developing local theories of response behaviour</td>
</tr>
<tr>
<td>Assessment findings</td>
<td>Illuminating the mechanisms and generative processes of test takers’ scores and assessment results</td>
</tr>
<tr>
<td>Methodological strategies</td>
<td>Micro-methodologies for test development and validation</td>
</tr>
<tr>
<td>Idiographic, qualitative data</td>
<td>Complementing nomothetic data with qualitative, idiographic data relevant to testing process and validity</td>
</tr>
<tr>
<td>Consequences of testing</td>
<td>Contributing to articulation of socio-culturally situated consequentialist frameworks in validity</td>
</tr>
</tbody>
</table>

**Confirmatory role.** This has been seen as the primary function of response processes research (Messick, 1989; Standards, 1999, 2014). According to the Standards (2014), validity evidence based on response processes refers to the “evidence concerning the fit between the construct and the detailed nature of performance or response actually engaged in by examinees” (p. 12). Under this approach, the role of response processes research is to establish whether the
observed response processes fit with the theoretically expected response processes defined by the construct theory.

**Exploratory role.** Investigating the fit between the observed responses and the theoretically expected response processes is not the only role that response processes can play in testing and validity. Collecting data about response processes in testing is not only a screening device for what does not fit, but it could also generate research findings valuable in their own standing. Investigating response processes that occur during testing can yield important novel results could go beyond the function of a screening device, such as: providing new information about the intra- and interpersonal dynamic of these response processes during testing, and understanding test takers’ subjective experiences and meaning making processes, even if some of these processes do not appear to neatly fit with the a priori theoretical framework of a construct. Moving from the status of screening tool to that of an independent research domain may offer response processes a more central role in psychometrics and validity, and may bring important substantive contributions to understanding and explaining the process of testing, as well as the validity of test scores interpretations. The fulfillment of this role would be possible only to the extent that the tight confirmatory frameworks operating in psychometrics and test validity can relax to make room for a more exploratory focus (Tukey, 1995).

**Response processes as psychometric parameters.** Although a somewhat daunting task, some response processes could potentially be mathematically formalized and included in designing and testing psychometric models of item responses. For example, Markus and Borsboom (2013) suggested that the psychometric modeling of item wording, interpretations or meanings would be a great stride forward in psychometrics. Zumbo et al. (2015) has also cited some research studies from the language testing domain that have addressed specifically this
issue (Fox, 2001, 2003, 2004). This role would be consistent with Embretson’s (2010) comment that: “The success of the construct modeling approach, especially for construct representation research, will depend on the ability of researchers and test developers to \textit{develop quantitative indices that define the theoretical mechanisms that are involved in the tasks}” (p. 195). This suggestion highlights the possibility that some response processes may play an important role in psychometric modelling or model based measurement, if the identified underlying mechanisms of test takers’ responses could be quantified and estimated via psychometric models.

Notwithstanding this aspiration, it is understandable that some of the response processes may never be and should never be amenable to a mathematical formalization. For example, it is hard to envision that intrinsically indeterminate response processes (e.g., meaning making processes) could be fully mathematically modelled. Nonetheless, whenever possible, it is worth exploring strategies of incorporating information provided by response processes research in the development of better calibrated psychometric models, even if not necessarily through mathematical formalism. Therefore, a critical question is whether, and to what extent could or should, response processes be quantitatively modeled within a psychometric model and at what level (e.g., intraindividual, contextual or sociocultural). In this sense, Molenaar’s (1985) discussion about idiographic filters, as well as Lamiell’s (1981) argument about idiothetic psychology offer some suggestions and invite researchers in the area of response processes to explore these possibilities.

\textbf{Response processes as feedback loops.} The findings about response processes may become an intrinsic part of test development and subsequent test revisions. Specifically, response processes can actively inform and shape the development, piloting, and refining of test items, and, thus, act as regulatory, feedback mechanisms in the process of test construction and test
revisions. This way response processes research may contribute to the development of more refined and better calibrated test items, either mathematically within a psychometric model (as discussed in the previous section) or as systematized qualitative findings.

The information provided by examining the response processes underlying answering test items may help: a) retain the items that engage test takers in responding in a manner congruent with that intended, b) refine the items that need alterations in order to reflect the desired response processes, or c) eliminate the items that solicit response processes that are not consistent with the response processes that the test intends to target. Specifically, in the case of the Rosenberg Self-Esteem Scale, response processes inquiry can provide information regarding what items engage most effectively test takers’ self-evaluation and self-regulation processes, and what items are less effective at mobilizing these response processes. For example, the findings of the empirical study presented in this dissertation suggest that the negatively worded items of the RSES are more effective than the positively worded items to engage test takers’ self-evaluative and self-regulation processes. Therefore, based on this information provided by the examination of the response processes underlying answering the RSES items, one may decide to keep and maybe add to the negatively worded items targeting different aspects of self-esteem.

**Response processes as research variables.** Response processes may play an important role as research variables that can be manipulated and investigated during experimental studies on process based validity (Bornstein, 2011). Specifically, researchers can first identify the core response processes that occur during testing, and then manipulate these under various experimental conditions in order to determine the relative importance and role of these response processes across various testing contexts. In addition, this data may also be used to refine the test items, and to accomplish a more accurate and flexible integration between the empirical findings
and the theoretical basis of the construct under study. If we consider response processes as research variables aptly to be experimentally manipulated, they may also become uniquely suited for investigating the response processes that occur over repeated testing occasions. This function of response processes would fit with the term “try and see psychometrics” that Markus and Borsboom (2013, p. 432) coined to describe an experimentally oriented psychometrics in contrast with the dominant “look and see” or correlational psychometrics.

**Response processes as theory building blocks.** The empirical findings on response processes connected with their respective theoretical rationales may function as local theory building blocks. Thus, studying these response processes could lead to substantive theory building, especially in the form of local theories relevant to contextualized testing and assessment practices. An example of a local theory could be a theory of response behaviour formulated for a certain test within a specific context (Borsboom et al., 2004). This way, the research on response processes may facilitate the transition from “grand theories” and "one size fits all" approaches towards local, flexible, and contextualized theoretical models and research heuristics.

As Freedman (2010) noted, robust scientific theories cannot be built solely on statistical inferences and on population level estimates. In order to witness any scientific breakthroughs, researchers also have to pay attention to ‘anomalies’ that usually go undetected during the statistical inference process but may be easily discovered during process observation activities. The impact of these observations highlights the role of response processes not only as substantive validity evidence, but also as potential building blocks in developing new theories.

**Response processes as testing findings.** Given the fact that response processes hold the promise to illuminate important mechanisms and generative processes of test takers’ scores, they
could form an intrinsic part of test findings, and of the interpretation/inferences based on these
test ratings. This function of response processes would fit well with process-based,
individualized, and collaborative assessment during which test takers’ response processes are
deemed essential for understanding the test findings and making individualized, collaborative

**Response processes as methodological strategies.** Response processes can be seen as a
sui generis way to develop, pilot, and implement process-based micro-methodologies for test
development and validation. For example, the investigation of the *imoments* in grief therapy
(Neimeyer, 2014) as well as the thin slice prediction method focused on noticing micro-level
response processes that may predict future disintegration/decompensation in clinical psychology
studies (Nalini & Rosenthal, 1992) are examples that illustrate how response processes may
become not just the object, but also the method, of investigation. Given their flexibility, response
processes can also integrate well with qualitative methodologies such as: conversation analysis
(Giorgi, 1981), polivocality, listening guide (Gilligan, 1982) or action method project (Young,
Valach & Domene, 2005).

**Response processes as subjective experiences, idiographic and qualitative data.**
Research on response processes in testing brings the focus in on individual and qualitative data
about test takers’ subjective experiences and processes that occur during testing. This type of
research data may contribute to a shift in the focus for validity and testing from abstract
constructs to the person who is taking a test. It is possible to imagine that, in this context, the
validity argument will not be exclusively the purview of test developers and test administrators
but it will also have to reflect the experiential findings of test takers’ response processes.
Moreover, response processes research may change the unit of data analysis from
interchangeable individuals aggregated across various samples to the person in context. This role of response processes would conjure the implementation of complementary, rich qualitative methodologies able to investigate subjective experiences and the meanings that emerge during the testing process. The rich, “thick descriptions” (Geertz, 2014) brought in via response processes may enrich the contemporary ‘thin’ psychometric and validity discourses, and promote the development of ‘thicker’ psychometric frameworks able to accommodate the richness of the new data.

**Response processes and consequences of testing.** Response processes research can become a fertile ground for articulating comprehensive consequentialist frameworks in the area of validity emphasizing social participation and responsibility as well as reflexivity in testing practice. Response processes research that takes into account test takers’ experiences, values, and the interactions with the larger socio-cultural context is uniquely suited to address the personal and social impact of test scores, and the consequences of testing in an integrated manner.
Chapter 5: Concluding Chapter

Introduction

The purpose of this dissertation project was to contribute theoretically and methodologically to research on response processes, which are considered one of the primary sources of validity evidence in testing (Messick, 1989; Standards, AERA, APA, & NCME, 2014). To this end, in chapter 1, I introduced readers to the topic of this dissertation by providing a critical review of the state of the art in response processes research; in chapters 2 and 3, I elaborated, empirically evaluated, and revised a theoretical model of response processes underlying responses to a self-report measure of self-esteem (i.e., the Rosenberg Self-Esteem Scale, Rosenberg, 1965), and, in chapter 4, I proposed several theoretical and methodological future research directions in response processes research. This concluding chapter will synthesize and discuss: a) the novel contributions of each of these chapters of the dissertation to the overarching topic of response processes research as they pertain to validity theory and validation practices, b) the limitations of this project, and c) the implications of this project findings for validity and testing.

Novel Contributions of the Dissertation Project

This dissertation project makes an important contribution to response processes research by elaborating and implementing a systematic, theoretically informed way of investigating response processes underlying answering self-report test items. In this sense, this dissertation represents a proof of concept regarding model building in response processes research, and its main focus was on developing a coherent strategy of iterative theoretical model building, providing empirical support for the model, and making suggestions for revising the model. Although limited to one self-report measure, this research project represents a significant methodological contribution given the scarcity of empirical investigations in response processes
research, the shortage of methodological guidelines about how to carry out this type of research, and the fact that most of the discussions regarding response processes have been largely theoretical with very little impact on validation practices. The model and the methodological framework for evaluating this model represent an exemplar for future empirical studies in this research domain, and an initial template that could be further elaborated and refined.

In addition, this dissertation project makes some theoretical contributions by proposing complementary epistemological and theoretical frameworks for studying response processes as well as future research directions in this domain, and has enlarged the validity base for a well-known self-esteem measure (the Rosenberg Self-Esteem Scale; RSES; Rosenberg, 1965), by providing validity evidence using response processes inquiry. Specifically, this dissertation is the first study that investigated the substantive validity evidence of the inferences made from the RSES scores using response processes inquiry. Moreover, some of the results of this investigation challenge some critical issues that have been raised in the research literature on the RSES (e.g., the wording or method effect, and, subsequently, the implications for the internal structure of the RSES).

Chapter 1: From limitations to possibilities. Based on the critical review of the theoretical and empirical state of the art in response processes research, the first chapter of this dissertation pointed towards the main limitations and shortcomings in the current research on response processes, such as: a) the lack of theoretical grounding and methodological guidelines in this research domain, b) the limited scope of response processes research (e.g., the exclusive focus on investigating cognitive response processes), and c) the scarcity of empirical investigations on response processes. Acknowledging these limitations represents an implicit suggestion about potential future directions in response processes research: a) developing
theoretical models and methodological frameworks that would allow a systematic investigation of response processes, b) expanding the scope of research studies to include non-cognitive and situated-contextualized response processes, and c) collecting empirical evidence based on response processes during validation studies as a requirement for providing adequate validity evidence.

**Chapter 2: Theoretically informed model based inquiry.** Chapter two made a significant contribution to the theoretical modeling of response processes underlying answering test items. Given the paucity of theoretical model building in response processes research, this represents a significant step forward towards elaborating a model-based approach in this research domain that could hopefully inspire a more systematic effort with respect to elaborating theoretical models of response processes underlying testing. Even beyond the specific response processes domain, the poverty of theoretically driven research, and of theory building and testing in contemporary social sciences in general is quite daunting. In this sense, the present study makes an important contribution towards a larger purpose than response processes seen as substantive validity evidence by encouraging a theory driven approach to empirical research. Moreover, to the extent possible, this research project attempted to maintain the focus of the inquiry on processes and not on contents or products. Considering the present significant conceptual and methodological limitations in process based models and theories in social sciences, although imperfect, this research project represents a step forward in this direction.

The investigation presented in chapter two made other important contributions by: (a) expanding the scope of response processes research beyond cognitive responses to self-referential processes, and (b) identifying the core self-processes relevant to answering self-report items. Given that self-reports are the most widely used strategies of collecting testing data in
social sciences, it is of particular importance to gain a better and more refined understanding of the response processes underlying answering self-report items. To my knowledge, the study presented in this dissertation is the first attempt to intentionally integrate relevant substantive theory in the domains of self, consciousness, and self-processes, on the one hand, with investigating response processes as substantive validity evidence for test scores obtained from self-reports, on the other hand. The investigation presented in this chapter materialized in the development of a theoretical model of responding to Rosenberg Self-Esteem Scale items, and an associated coding scheme that could be further elaborated and concretized in future investigations examining the response processes underlying answering self-report items.

Chapter 3: Validity evidence for the Rosenberg Self-Esteem Scale. Chapter three of this dissertation accomplished a dual objective and, thus, made a twofold contribution by: a) empirically evaluating and revising the theoretical response processes model of self-esteem, and b) providing validity evidence for the interpretations made of the RSES scores using response processes methodology. Although the RSES is the most widely used self-report measure of global self-esteem, no research studies have investigated the substantive validity evidence for the RSES scores using response processes inquiry before this current project. Thus, the findings of this chapter represent a significant contribution to the validity of the inferences made from the test scores obtained from the RSES.

Chapter 4: Looking towards the future of response processes research. Chapter four represents a theoretical contribution to the conceptualization of response processes with respect to potential future research directions in this domain. The chapter provided a critical evaluation of shortcomings of response processes with an eye to addressing some of these in future directions: (a) expanding the theoretical and methodological ground of response processes
research by exploring paradigmatic alliances, alternative epistemologies, and complementary research methods, b) expanding the scope of response processes research from studying individual cognitive processes to investigating interpersonal, situated, and sociocultural response processes, and c) redefining the roles of response processes in validity research.

Limitations of the Presented Research

Given the novel approach of this dissertation project, limitations are not only unavoidable but they represent important lessons and critical feedback to inform future similar research projects. The next paragraphs will address theoretical, methodological and practical limitations together with some suggestions on how to address these shortcomings in future studies.

Theoretical limitations. Rogosa (1987) wrote that “building scientific models for the social sciences is very hard work and requires orders of magnitude more thought, preliminary empirical research, careful data analysis, and creativity in statistical modeling than is now evident” (p. 365). This quotation summarizes very pertinently the intrinsic challenges of elaborating and testing theoretically driven response processes models. Building these models requires in-depth and very well differentiated substantive knowledge, deciding what knowledge is relevant for what purpose, and determining the appropriate level to model that knowledge (e.g., individual, social, contextual), among other requirements. Building such models almost goes against the grain of the current research practices in social sciences, and against many assumptions and practices in psychometrics and validation practices.

Specifically, one of the limitations that I faced in my attempt to build a response process model was the perplexing ‘thinness’ and the overwhelming paucity of process based models and theories in psychology and social sciences. This is not just a limitation of a research study or of response processes research but this represents a built-in limitation of several disciplines, and,
quite likely, of a scientific paradigm (Freedman, 2010; Kazdin, 2007; Toomela, 2009), which, in turn, may reflect the unfortunate effect of too many decades of dustbowl empiricism in an academic tradition largely devoid of interest in teaching students the bases of the philosophy of science, epistemology, and theorizing.

This limitation has significantly impacted the theorizing efforts presented in this study. For instance, the most widely used measure of self-esteem, the Rosenberg Self-Esteem Scale, does not have any validated theoretical model of the construct that it purports to measure. Moreover, before this study, there were no investigations about the response processes underlying responding to the RSES items, in spite of this scale having been widely used in research, program evaluations, and clinical settings as the ‘gold standard’ measure of self-esteem. To me, this is perplexing, and, based on the data from several reviews, it is very likely that most of the extensively used self-report measures in psychology have never been examined with respect to the response processes underlying answering their items.

The substantive knowledge required to develop a response process model seemed largely divorced from the psychometric models and principles, and investigating qualitative response processes fit quite awkwardly with the internalized pressure towards statistical inferences and mathematical formalism. All of the philosophical discussions about the role of response processes models in validity are not only abstract but also largely unhelpful in building up on them. The more specific conceptual scaffolding for building and testing such a model was missing. As a result of this, the initially proposed model was not only generic but it was developed at too abstract and general a level. Hence, it required revisions mostly in the sense of increasing its specificity. One of the lessons learned in this process is that it is critical to correctly tailor the model in terms of specificity and the level of interest (individual or trans-individual
response processes) as well as specifying the interactions and mutual influences among the many levels. Moreover, I think a model should be specific enough to be amenable to experimental manipulations of at least some of its components. Therefore, it is very difficult to propose a generic model of response processes, and each model should be tailored not only to the construct targeted by the test but to the specific testing situation, context and purpose.

**Methodological limitations.** Given that the empirical data analyzed and discussed in this dissertation were collected via a think aloud protocol and cognitive interviewing, the situated and social response processes emerged solely via participants’ spontaneous self-reporting, because the implemented method was not the most conducive to a direct elicitation and observation of these processes. In the future, it will be important to use other methods and settings that could be more conducive to the elaboration of a socially situated response processes model. Notwithstanding the clear limitations of the one to one cognitive interviewing method, most participants in this study spontaneously reported situated responses, such as: social desirability, impression management, and how social meanings shaped their responses to the RSES items.

Although not presented in this dissertation, listening to the audio-recordings from the interviews revealed a wealth of micro-processes that reflected the dynamic of how participants negotiated across multiple meaning systems to come up with a decision about the RSES item ratings. For future studies, it would be important to pay more attention to voice analysis and to other non-verbal markers (e.g., video-recordings) that are indicative of dynamic processes underlying responding to test items. I found transcript analysis quite dry and devitalized in comparison to the audio-recordings in terms of capturing the dynamic aspects of response processes. Therefore, it may be possible that the voice could be a better way to investigate the dynamic aspects of response processes.
Another important observation that I made while listening to recordings was the fact that one has to carefully and mindfully tune in to the dynamic processes and to intentionally bypass content that, at least during the first listening, definitely take centre stage. My experience was that it was extremely easy to get lost in content pretty much the same way that beginner therapists get lost in clients’ stories. And once one is lost, it is very hard to find the path back. It took me several attempts to learn how to listen for the flow and the process and to sustain that focus.

This difficulty was also encountered by participants during the TAP. Hence, it will be important that future research studies train respondents to tune into their mental processes and not just mental contents. To this end, methods such as focusing and mindful responding may be helpful to use before and during collecting information via a TAP.

Another suggestion for future studies would be to directly address the social and contextual dimensions of the testing situation. A very simple question that, in retrospect, I regret that I did not ask during cognitive interviewing would have been: “How is it for you to share this information with me in this situation?” In addition, integrating experimental manipulation and implicit and physiological measures with interviewing techniques could also be useful in future studies on response processes.

**Sample limitations.** The limited sample size and the highly exploratory nature of this study caution against any attempts to generalize these findings beyond this sample. However, the main point of this type of investigation is not statistical generalization but rather to gain an in depth understanding of response processes engaged in when answering self-report items. Nonetheless, more similar and hopefully better designed studies of this kind are necessary to draw any firm conclusions.
A unique characteristic of this particular sample was the unexpected diversity of the sample. Although this considerably enriched the findings of this study, it also cautions against interpreting the proposed response processes model as cross-culturally valid. Future research studies should be designed and implemented with participants from different cultural backgrounds to test and refine this proposed model.

**Implications of the Presented Findings**

I will start this brief discussion about implications with a question that begs to be asked at the end of this dissertation project: given all this work, should we devote any more effort and time to studying response processes as substantive validity evidence? Is this realistic? What good could ever come out from this?

My unedited answer to the question of whether we should study response processes is absolutely “yes”: if we are truly serious about finding out what we are doing or measuring, if we really care about ethical testing and social consequences, and if we want to do more than crunch numbers in statistical models and then present these results as validity evidence, this is an unescapable and necessary route. If we truly want to improve our measures, and our testing practices, we must look at response processes. It is important to note and repeat, as many times as necessary, that no statistical analysis, no matter how sophisticated, can substitute or circumvent systematic observations and theory building in testing and validity. Response processes research represents a critical piece of this endeavour.

**Socially situated constructs.** Research on response processes brings to the fore the role of socially constructed response processes that need to be taken into consideration in making inferences and interpretations about test scores across different testing situations and social contexts. In this way, response processes research may contribute to enriching the grounds and
scope of consequential validity (Messick, 1989; Hubley & Zumbo, 2011), and may also contribute to expanding the focus in validity discourse from theoretical, abstract constructs towards historically situated, value driven, and socially responsible practices of testing and assessment which necessarily require one to take into consideration real persons in their life contexts for whom testing has palpable and sometimes life altering consequences. Under these circumstances, validation efforts should focus on validating situated constructs that emerge at the intersection between test items and test taker. Therefore, elaborating the interface between abstract construct theory and the praxis of testing becomes a pressing necessity.

**Person focused response processes models.** The shift from abstract and isolated constructs to the person who is taking the test in a contextualized interaction with the construct targeted by a test brings up a new model of imagining testing: the Person-Test Item-Person-Test Score model versus the Test Item-Person-Test Score model or even the most simplistic Test Item-Test Score model. This new model brings to the fore the study of the intentionality and purpose that bridge items and responses formulated within a situated socio-cultural and linguistic context. Before the Test Item, there is a Person who is shaping the meaning of the test items and the testing situation, and between the Test Item and Test Score there is a space where the response processes of test ratings are generated, a space filled with intentionality, meaning, and purpose, not just cognitive mechanisms.

**Experiential sources of validity.** This approach would recognize that the test taker is contributing to the meaning of the construct as much as the test developer (i.e., test taker based validity or experiential validity). Answering the question: “what does the construct mean for the test taker?” is as important as knowing what is the theory of the construct. This does not exclude the existence of shared meanings of the construct or the possibility of investigating the expected
response processes based on the construct theory. It is very likely that, because we are part of meaning making communities, we share meanings and there are few highly idiosyncratic or unique meanings. However, even shared meanings may be mobilized differently in different contexts by different test takers.

**Contextualized validation frameworks.** In the latest edition of the *Standards* (2014), context is explicitly seen as a potential threat to test fairness, and, implicitly, as a threat to the validity of inferences based on test scores. Nonetheless, the meaning of any test scores, like any meaning, makes no sense in the absence of context (Leontiev, 1985). Proposing a definition of validity focused on the meaning of test scores, and, at the same time, reducing test context to a potential error makes no sense epistemologically. Any test score is produced within a context, and context is an intrinsic part of the test scores. Therefore, any validity discourse should be contextualized and situated.

It is important to be more rigorous with respect to what we mean by ‘context’ given that contextualism tends to be a trendy umbrella term that may conceal many different meanings ranging from radical relativism to simple qualifiers of invariant findings. Therefore, it matters to distinguish among several levels of “contextualization” in validity discourse. In this sense, the conceptual framework proposed by Longino’s (1990) contextual empiricism may be useful. In her theorizing about context, Longino distinguishes among three layers of contextualism: (a) epistemic contextualism that refers to epistemic justification in the sense that any knowledge is relative to background assumptions and such assumptions are needed to establish the relevance of empirical evidence to a hypothesis or a theory, (b) community-level contextualism that claims that the objectivity of scientific knowledge is a function of a community’s practice rather than an individual scientist’s observations and reasoning, and (c) sociocultural and historical-axiological
contextualism that emphasizes the role of values in science and how values belonging to the social and cultural context of science can enter into epistemic justification via background assumptions. With respect to validity, epistemic contextualism may refer to the context of justification and explanation of test scores depending on epistemological assumptions whereas community and socio-cultural contextualism emphasize the axiological and consequential aspects of validity together with the focus on socially situated constructs and the lived experiences of test takers.

**Hermeneutic-interpretive focus.** Research on response processes can complement and enrich a psychometric tradition exclusively concerned with objective measurement and quantification at the expense of investigating test takers' subjective experiences and the role of interpretations and meaning-making processes during testing. There are several implications of shifting towards an interpretive focus in validity discourse: (a) de-emphasizing theoretical relics about reified constructs and, consequently, about the test scores reflecting these constructs invested with unequivocal, a priori meanings; (b) integrating subjective meaning and idiographic dimensions in test construction and validity, and (c) accounting for interpretative, hermeneutic aspects in psychometrics and validity (Markus & Borsboom, 2013). Interpretative, hermeneutic aspects are consistent with Messick’s view on validity (Messick, 1989), and with the *Standards* definition of validity as truthfulness and meaningfulness of test scores. This view is also consistent with Markus and Borsboom’s observation that meaning making processes are most notably underdeveloped in test validity, and that important strides in psychometrics can be made by refocusing on researching these aspects. Response processes may be ‘the royal way’ to explore these meanings. Also, this view may be consistent with the definition of validity from
the *Standards* (1999, 2014) that posits that validity is mainly about the interpretations of test scores and finding the appropriate empirical and theoretical evidence for that.

‘Thick Psychometrics’. An alternative approach to some of the more traditional psychometric practices would aim to explicitly, systematically and rigorously integrate contextual sensitivity and meaning making/interpretive processes into psychometric theory and models, test development and validation, and during the testing process, either in formal mathematical ways or informally (Markus & Borsboom, 2013; McGrath & Johnson, 2003). This integration would not imply a cumbersome, automatic, and meaningless quantification of every single testing aspect that might influence the inferences made from the test scores but it should render a thoughtful and systematic understanding of the response processes that are critically involved and relevant for answering a specific test. This understanding would lead to an elaboration of integrative methodological frameworks to accommodate both the quantitative and qualitative aspects of these response processes.

In other words, this approach involves a trained capacity to discern the response processes that could contribute to an improved testing process from those which may be redundant for the purpose of that testing context. To use a metaphor from nursing qualitative research that incorporated the distinction of "thin" versus "thick" descriptions (Dueck & Reimer, 2003), this alternative approach would be a way of developing a “thick psychometrics” (McDonald, 2012, unpublished manuscript). The purpose would be to move from “thin”, abstract, isolated constructs towards ‘concrete generalizations’ (Piaget, 1981; Vygotski, 1996), situated constructs, emergent meanings, and towards richer assessment and psychometrics models that take into account test takers' generative and contextualized response processes.
This goal closely mirrors a tendency noticeable in some more recent areas of psychometrics where the newest statistical models have become more open towards acknowledging and accounting for context and intra-individual variability using quantitative frameworks (e.g., latent class modeling, mixture growth modeling, and multilevel validation practices). It also fits with the model based measurement approach that allows for formalizing some response processes to become an intrinsic part of the psychometric models (Embretson, 2010). It is very likely that not all qualitative, idiographic information pertaining to generative response processes relevant to testing and validity can or should be modelled mathematically via mathematical formalism. Resorting to math as a band aid is unlikely to help. Thus, comprehensive and flexible frameworks that allow an easy synthesis of the quantitative and qualitative information in response processes research need to be developed and implemented.
References


Rethinking social inquiry: Diverse tools, shared standards (pp. 207–219).

Lanham, MD: Rowman & Littlefield.


Psychological Review, 111, 1061-1071.


University Press.


Psychology, 4, 275-296.

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of
assessment: Complementary paradigms. Psychological Assessment, 9, 374-385.


Assessment, 74, 2-14.

Slife, B. D., Williams, R. N. & Barlow, S. H. (Eds.), Critical issues in psychotherapy:

Psychology, 43, 93-105.

Journal of Personality Assessment, 82, 35-38.

techniques: A life-world approach. In M. Hersen, M. J. Hilsenroth, & D. L. Segal (Eds.),
Comprehensive handbook of psychological assessment: Vol. 2. Personality assessment

NY: Other Press.


doi:10.1177/0146167203259930


36. 276–289.


McDonald, M. J. (2012). Thickening the thin discourses of psychometrics. *Unpublished manuscript*. Trinity Western University, Langley, BC.


Molenaar, P. C. M., & Valsiner, J. (2005). How generalization works through the single case: A


Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J.


    Valsiner, P. Molenaar, N. Chaudhary, and M. Lyra (Eds.), *Handbook of dynamic process
    methodology in the social and developmental sciences* (pp. 99-121). New York:
    Springer.

Wang, X., & Sireci, S.G. (2013). Evaluating the cognitive levels measured by test items using
    item response time. *Paper presented at the annual conference of the American
    Educational Research Association, San Francisco, USA.*


    neuroscience and psychiatry.* Cambridge, UK: University Press.

    systems perspective. *Journal of Personality, 70,* 851–900. doi.org/10.1111/1467-
    6494.05026

    doi:10.4324/9780203587874

Zumbo, B. D. (2007a). Three generations of differential item functioning (DIF) analyses:
    Considering where it has been, where it is now, and where it is going. *Language
    Assessment Quarterly, 4,* 223–233.

    & S. Sinharay (Eds.), *Handbook of statistics, Vol. 26: Psychometrics* (pp. 45–79).

Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications


Appendix A: Poster for Participants’ Recruitment

The University of British Columbia
Dept. of Educational and Counselling Psychology, and Special Education
2125 Main Mall, Vancouver, BC, Canada V6T 1Z4
Tel: (XXX) XXX-XXXX Fax: (XXX) XXX-XXXX

ADULTS (19+) NEEDED FOR STUDY ON TEAMWORK:
HOW PEOPLE TALK AND PLAY WITH EACH OTHER

Men and women ages 19 years and older are needed to take part in an exciting research project. Dr. Anita Hubley, a professor at the University of British Columbia (UBC) and Mihaela Launeanu, PhD student at UBC, are conducting a study about teamwork and how people talk and play with each other during various social interactions.

In this study, you will need to complete several self-report questionnaires about different personal characteristics, a demographic sheet with information such as gender, age, ethnicity etc, and to take part in a fun experiment about playing together with others. This study will take approximately 2 hours to complete in one single session.

If you have any questions about the study or would like to take part, please call and leave a message for Mihaela Launeanu at the Adult Development and Psychometrics Lab at UBC at XXX-XXX-XXXX.

Thank you!
Appendix B: Informed Consent

The University of British Columbia

Dept. of Educational and Counselling Psychology, and Special Education

2125 Main Mall, Vancouver, BC, Canada V6T 1Z4

Tel: XXX-XXX-XXXX Fax: XXX-XXX-XXXX

Informed Consent

The purpose of this Informed Consent is to ensure that you understand the purpose of this study, what you are being asked to do, and your rights as a participant.

Study of Teamwork: How People Talk and Play with Each Other

Principal Investigator. Dr. Anita Hubley, Professor in the Department of Educational and Counselling Psychology and Special Education at the University of British Columbia, is the principal investigator.

Co-Investigator. Mihaela Launeanu, PhD student in the Department of Educational and Counselling Psychology and Special Education at the University of British Columbia.

Purpose. This study, which is being conducted for Mihaela Launeanu’s doctoral dissertation, will help us understand better various strategies that people use when they work or play together. This study will also examine how people feel about themselves and how they deal with different situations that may occur as part of their working or playing together.

What is required. You will be asked to complete some questionnaires asking about various personal characteristics such as self-esteem and meaning in life, and a demographic sheet. During completing some of those measures the investigator will engage you in a conversation about how you felt or thought as you answered those items. There will be no right or wrong answer but rather an open conversation about how was it for you to answer to those questions. This will provide some indication about how you can openly communicate your thoughts and attitudes during a brief interpersonal interaction. Then, you will take part in a fun social activity that involves a ball tossing game between three people. Depending on which group you will be assigned to you may take part in a fun social game that will require you to play a ball tossing game with two other participants.
**How long does it take?** This study will take place in one single session and the entire session is expected to take approximately 2 hours.

**Potential risks.** Usually, people taking part in this type of research studies do not experience any harmful consequences during or after the study. However, in the case that you will experience any negative emotions during this study I encourage you to report those to the researcher as soon as possible so that your discomfort will be addresses promptly and effectively. In the case that, after you finish the study and/or after you leave the research lab, you will experience any negative effects potentially connected with the activities included in this study you will be given a list of resources that you may contact for further assistance.

**Monetary Compensation.** Participants in this study will not receive monetary compensation for their time, but they may be reimbursed for the typical cost of parking while taking part in the study.

**Anonymity/Confidentiality.** Your name will not be recorded on any of your questionnaires. Instead, you will be assigned an identification number that will appear on each of your questionnaires. The questionnaires will be kept in a separate place from the informed consent forms. All study materials are kept in a locked room at UBC.

**Important information.** If you have any questions or would like further information about this study, please contact Mihaela Launeanu at XXX-XXX-XXXX (Lab) or Dr. Anita Hubley at XXX-XXX-XXXX (office). If you have any concerns about your rights or your treatment as a research participant, please contact the Director of the UBC Office of Research Services and Administration at XXX-XXX-XXXX.

Participant’s Signature:

Date:
Appendix C: Personal Demographic Form

*Please answer the following demographic questions as accurately as you can. All the information that you provide on this form is confidential. This information is collected for the purpose of describing the study sample.*

1. Age: _____ years  
2. Date of Birth: _____________ ___ 19___  
   (month) (date)

3. Are you:  
   □ male  
   □ female  
   □ other

4. What is the highest level of education that you have?  
   □ 0-8 years schooling  
   □ Some high school (no diploma)  
   □ High school graduate (received diploma or G.E.D.)  
   □ Some college/university, but no degree  
   □ College/university graduate (received degree)  
   □ Graduate degree completed  
   □ Other (please specify): ______________________________

5. What is your primary ethnic/racial/cultural background? (please check one box only)  
   □ Aboriginal / First Nations (includes Métis, Inuit)  
   □ African (e.g., South Africa, Nigeria, Kenya)  
   □ East Asian (e.g., Chinese, Japanese, Korean, Taiwanese)
☐ Hispanic, Latin American (e.g., Latino, Mexican)

☐ Pacific Islander (e.g., Australia, New Zealand, New Guinea)

☐ South Asian (e.g., Indonesian, East Indian, Philipino)

☐ Southeast Asian (e.g., Cambodian, Filipino, Vietnamese)

☐ West Asian (e.g., Afghan, Arab, Iranian, Iraqi, Turkish)

☐ White (e.g., Caucasian, Anglo, European origin)

☐ Other (please specify): ____________________________

(Note: Please do not record nationality, such as “Canadian”, for this question.)

7. Today’s date is: __________, __________ ______ 2011

   (day)          (month)         (date)
Appendix D: Sample of Possible Probes for Cognitive Interviewing

All self-report items will be presented one by one on separate cards so that participants can focus their attention on each item separately.

a) Instructions for Think Aloud Protocol (TAP)

“During the next activity I will ask you to tell me what you are thinking or feeling when you are answering each question that I will show you. As much as you can, just say loudly, as detailed as possible, everything that goes on through your mind as you read and answer this item. Please do not worry about whether what you say makes sense or if it is expressed in coherent or complete sentences. You could just say words or group of words, images, metaphors, anything that may convey your reactions, feelings and thoughts that you experience as you answer each question. The point of this activity is to allow yourself to say everything that you are aware of as you answer each item”

Before each item I will say the following instructions and I will encourage participants’ free, spontaneous expression of thoughts/feelings/reactions:

“Tell me what are you thinking or feeling while you are answering this question”

“Be aware and tell me how you are coming up with your answer to this question”

b) Possible Questions for Verbal Probing

“Now I will ask you some more specific questions about how you answered this item. I would like you to take your time and go back to the moment when you first read the item.”

“What was your first impression/reaction/thought when you read the question?”

“Have you answered this question or a similar question before?”

“If yes, did you answer now the same way that you answered it before?”

“Were you influenced by your previous answer?”

“What feelings/memories does this question evoke in you?”

“If you were to restate this question into your own words what would you say?”

“How did you interpret the word X within that question?”

“Did you have a reaction to certain words such as ‘useless’, ‘failure’; ‘worth’?

“Did those words influence your reactions and answering the question more than others?”

“How did you decide how much you agree or disagree with this item?”
“Did you decide how much you agreed or how much you disagreed or both?”

“Was it different for you to decide if you agreed versus to decide if you disagreed?”

“Did you interpret/perceive the suggested response options (‘strongly agree’, ‘agree’, ‘disagree’, ‘strongly disagree’) as related to each other? As degrees of agreement/disagreement?”

“Do you think that you would have responded differently if instead of ‘strongly agree’, ‘agree’, ‘disagree’, and ‘strongly disagree’ you were asked to rate your answer on a scale from 1 to 4 (1= strongly agree and 4= strongly disagree)? In what way?”

“What does agreement mean to you in the context of this question?”

“What role did qualifiers such as ‘at times’, ‘on the whole’, ‘sometimes’ play in your decision for one answer or the other?”

“What information did you take into consideration when deciding how much you agree with each statement (e.g., previous experiences, other people’s feedback, expectations, something good happened to me…”

“How did you make your decision regarding answering this item” or “What did you take into consideration when you answered this item”?

“Did other/previous items influence your decision on this particular item?”

“Does this answer represent your current state or do you see it as a more permanent trait of yours”?

“Was it easy/straightforward to answer this item; what was easy, what was hard/unclear/confusing?”

“How important/relevant is the characteristic described by this item for you in your daily life?”
Appendix E: Mental Health Resource Sheet

24-Hour Resources: 911

For emergencies, please call 911.

Crisis Centre: 1-800-784-2433

http://www.crisiscentre.bc.ca/get-help/

The Crisis Centre provides free confidential and nonjudgmental emotional support 24 hours a day, 7 days a week for individuals experiencing feelings of distress or despair.

Vancouver General Hospital Emergency Department: 604-875-4111

VGH provides inpatient and outpatient treatment for mental health difficulties.

Fraser Health Emergency Mental Health Services: 877-384-8062

Mental Health Services provides clinical and emotional support to individuals in crisis.

Regular Business Hour Resources

Vancouver Community Mental Health Services (VCMHS)

VCMHS offers diagnosis, treatment, individual and group therapy, rehabilitation, and consultation for children, youth, and adults/older adults with persistent mental health issues.

Below are the contact numbers by geographic area:

Grandview/Woodlands 604-251-2264

Kitsilano/Fairview 604-736-2881
Midtown  604-872-8441
Northeast  604-253-5353
South  604-266-6124
Strathcona  604-253-4401
West End  604-687-7994
West Side  604-873-6733

Fraser Health Mental Health (FHMH) Care Programs

FHMH programs offer assessment, treatment, crisis intervention, and community support to youth, and adults/older adults with mental health issues. Below are the contact numbers by geographic area:

Abbotsford  604-870-7800
Agassiz  604-793-7160
Burnaby Central:  604-453-1900
Burnaby North:  604-949-7730
Burnaby South:  604-777-6870
Chilliwack  604-702-4860
Delta 56th St:  604-948-7010
Delta North:  604-592-3700
Hope 604-860-7733
Langley 604-514-7940
Maple Ridge 604-476-7165
Mission 604-814-5600
New Westminster 604-777-6800
Surrey 604-953-4900
Tri Cities 604-777-8400
White Rock 604-541-6844

UBC Student Only Resources

Counselling Services: (604) 822-3811

http://www.students.ubc.ca/counselling/

Free individual therapy and low-cost group therapy is provided for UBC students at UBC Counselling Services, which is located in Brock Hall. Registration for drop-in hours is available each weekday beginning at 10 am, but it is recommended that students call to make an appointment.

Student Health Services: (604) 822-7011

http://www.students.ubc.ca/health/service.cfm

The Student Health Services provides free psychiatric consultation to UBC students, but note that there is often a waitlist.
Speakeasy Peer Support, Information, and Referral: (604) 822-9246 or e-mail: speak@ams.ubc.ca

http://www2.ams.ubc.ca/index.php/services/category/speakeasy/

Speakeasy provides drop-in support and peer-referrals for students. Located in the SUB, this program is only available during the school year (Sept.-April).

**Online Resources**

**Crisis Centre**

http://www.crisiscentre.bc.ca/get-help/

This website provides more information about the Distress Line, the YouthBC online chat service, and other resources for those in crisis.
**Appendix F: Rosenberg Self-Esteem Scale (Rosenberg, 1965)**

**Instructions:** Below is a list of statements dealing with your general feelings about yourself. If you strongly agree, circle **SA**. If you agree with the statement, circle **A**. If you disagree, circle **D**. If you strongly disagree, circle **SD**.

<table>
<thead>
<tr>
<th></th>
<th>statement</th>
<th>SA</th>
<th>A</th>
<th>D</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>On the whole, I am satisfied with myself.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>2</td>
<td>At times, I think I am no good at all.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>3</td>
<td>I feel that I have a number of good qualities.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>4</td>
<td>I am able to do things as well as most other people.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>5</td>
<td>I feel I do not have much to be proud of.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>6</td>
<td>I certainly feel useless at times.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>7</td>
<td>I feel that I’m a person of worth, at least on an equal plane with others.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>8</td>
<td>I wish I could have more respect for myself.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>9</td>
<td>All in all, I am inclined to feel that I am a failure.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>10</td>
<td>I take a positive attitude toward myself.</td>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>SD</td>
</tr>
</tbody>
</table>
Appendix G: Theoretical Coding Scheme for Response Processes on the Rosenberg Self-Esteem Scale

Process Coding (PC)

- **Self-processes** (code PC 1)
  - *Self-referenced responding* (code PC 11)
    - Explicitly refers to self (code PC 111)
    - Explicit use of first person reporting (code PC 112)
    - Spontaneously distinguishes self from others (code PC 113)
  - *Self-evaluation processes* (code PC 12)
    - Lateral comparisons (code PC 121)
    - Upward comparison (code PC 122)
    - Downward comparisons (code PC 123)
    - Past self-comparisons (code PC 124)
    - Future self-comparisons (code PC 125)
  - *Self-regulation processes* (code PC 13)
    - Enhance (code PC 131)
    - Maintain (code PC 132)
    - Protect (code PC 133)
    - Restore (code PC 134)
    - Compensate (code PC 135)
  - *Self-attributions* (code PC 14)
    - Internal and stable (code PC 141)
    - External and temporary (code PC 142)
  - *Self-expression processes* (code PC 15)
- Social desirability (code PC 151)
- Semantic responding (code PC 152)
- Other impression management responses (code PC 153)

- **Self-related cognitive processes** (code PC 2)
  - **Autobiographical memory processes** (PC 21)
  - Elaborating autobiographical narrative about self-esteem (code PC 211)
  - Accessing episodic events related to self-esteem (code PC 212)
  - Accessing “self-esteem moments” (code PC 213)
    - Situation of past failure (code PC 2131)
    - Situation of past threat to self and self-esteem (code PC 2132)
    - Past and evoked situation of rejection (code PC 2133)
    - Death or death anxiety (code PC 2134)
    - Anticipated situation of failure or rejection (code PC 2135)
    - Anticipation of a challenging situations (code PC 2136)
    - Imagining a significant, valuable situation in present or future (code PC 2137)
  - Accessing declarative (semantic) self-knowledge (code PC 22)

- **Self-related emotional processes** (code PC 3)
  - Reflective valence (code PC 31)
    - Positive valence (code PC 311)
    - Negative valence (code PC 312)
  - Self-conscious emotions (PC32)
  - Basic emotions (PC33)

- **Attitudinal processes** (code PC 4)
- Cognitive elaboration (code PC 41)
- Cognitive consistency (code PC 42)
- Propositional evaluation (code PC 43)
- Behavioural activations (code PC 44)

**Content coding (CC)**

- **Self-representations** (code CC 1)
  - competence (CC 11)
  - self-worth (CC 12)
  - likability (CC 13)
  - expectations (CC 14)
  - role identifications (CC 15)
  - needs and motives (CC 16)

- **Self-narratives** (code CC 2)
  - competence (CC 21)
  - self-worth (CC 22)
  - likability (CC 23)
  - expectations (CC 24)
  - role identifications (CC 25)
  - needs and motives (CC 26)

- **Self-beliefs** (code CC 3)

- **Implicit theories of self** (code CC 4)

- **Sources of self-esteem** (code CC 5)
  - Relationships (CC 51)
- Work (CC 52)
- Achievements (CC 53)
- Feedback (CC 54)
- Other self-esteem sources (CC 55)

  **Interactionist, situated response processes (IC)**
  - Association dimension (code IC1)
  - Dominance dimension (code IC2)
  - Intimacy-formality dimension (code IC3)

  **Socio-cultural processes and meanings (SC)**
  - Social meanings (code SC1)
  - Language (code SC2)
  - Culture (code SC3)
Appendix H: Response Processes Underlying Responding to RSES: Item Level Analysis

**Item 1: On the whole, I am satisfied with myself.**

<table>
<thead>
<tr>
<th>Response Processes</th>
<th>Verbatim</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluative processes:</strong></td>
<td></td>
</tr>
<tr>
<td>Meeting expectations</td>
<td>“I fell short on my own expectations many times”, “To be satisfied means to have high expectations about how happy you are with yourself but you also have to make sure that you are satisfied with the others and that others are satisfied with you”</td>
</tr>
<tr>
<td>Self-improvement</td>
<td></td>
</tr>
<tr>
<td>Downward social comparisons</td>
<td>“I agree…the reason is I see people around me and I see more less fortunate people around me and I have that thing, because the place that I’m living here there is a café nearby and whenever I go there, there are, there are mentally retarded people who come and handicapped people who come and all people who are lonely and all those people they come there. And I always see them and I feel: oh my life is so lucky…I am satisfied with myself”</td>
</tr>
<tr>
<td>Upward social comparisons</td>
<td>“I am not yet as good as others in many ways”, “I can see how strong, how good other people are and, yes I compare that with my situation, and I am not that satisfied, I have lots to improve”</td>
</tr>
<tr>
<td>Temporal comparisons</td>
<td>“Disagree…I don’t feel that way right now…right at the moment, but…but I’m answering it that way simply because of...of how I feel about the past, and it is present in my mind, because I’m learning from it…I’m growing from it, so it’s there”; “…a lot of things I can improve and especially when I was young, I’m not that satisfied with myself so as I am getting mature, I’m trying to improve myself”</td>
</tr>
<tr>
<td>Comparing with ideal/desirable self</td>
<td>“I feel that I could become a better self. I have to. I want more from myself. I want to be better in the future. So I disagree”</td>
</tr>
<tr>
<td><strong>Self-regulation processes:</strong></td>
<td></td>
</tr>
<tr>
<td>Positive self-talk</td>
<td>“I stand by myself and say to me ‘you did that, you succeeded, well done’. So I am satisfied with that.”</td>
</tr>
<tr>
<td>Response Processes</td>
<td>Verbatim</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Using feedback from others</td>
<td>“How others see me is important…I am satisfied with myself when others like me and see me as a good person”</td>
</tr>
<tr>
<td>Self-related emotionality</td>
<td>Gut feeling</td>
</tr>
<tr>
<td></td>
<td>“A gut feeling...yeah, yeah. I'm satisfied with myself. I could be doing better in some areas of my life, but I also feel pretty proud of my accomplishments, and...yes, overall I’m satisfied, it’s a gut feeling”; “I don’t know how to explain this. I am not thinking about it, I just feel satisfied. It is a feeling. I don’t know how to explain it more. I feel it here, right in my heart: I am satisfied with myself”;</td>
</tr>
<tr>
<td></td>
<td>Self-liking</td>
</tr>
<tr>
<td></td>
<td>“I really like myself...a lot”, “I like myself and feel comfortable with myself”, “It may sound conceited but I do like myself, I am satisfied with who I am”</td>
</tr>
<tr>
<td></td>
<td>Self-reflective moral agency</td>
</tr>
<tr>
<td></td>
<td>“Everything is good with my moral compass, I am satisfied with myself”, “There’s no conflict with my moral values, I am at peace and satisfied”, “I am satisfied because I made the right choices in life”, “I choose to act in agreement with my moral values”</td>
</tr>
<tr>
<td>Cognitive processes</td>
<td>Counterfactual thinking</td>
</tr>
<tr>
<td></td>
<td>“If I were not satisfied, I would not enjoy spending time with myself, I would not be proud of my accomplishments, of my work, of my family. So, yeah, I guess I must be satisfied if I have all these”</td>
</tr>
</tbody>
</table>
**Item 2: At times, I think I am no good at all.**

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong> Temporal comparisons</td>
<td>“I'm going to again go with an 'agree' there, mainly because of what my past and what I've come back from. So, I was quite low at one point in time. I was quite depressed...quite upset at a lot of things, and I have been going back in the right direction for some time now, but I haven't made it all the way back, so...I'm going to put a 3 there”</td>
</tr>
<tr>
<td><strong>Social comparisons</strong></td>
<td>“Well here you are always competing with people. And sometimes you are above, and sometimes you are on the bottom and, and in those time I, I say no, I am not good at all, sometimes you discourage yourself and then you realize that you, well at that time you, you were at the bottom. In comparisons with others, always”</td>
</tr>
<tr>
<td><strong>Meeting expectations</strong></td>
<td>“There are times when I think I’m not good enough as to what I feel I should be. I was doing my best at what I thought was a good thing but that was not enough for specific people and the same goes the other way. So those were the times when I think I’m not good and those were the times that triggered me so I should improve on this thing”; “I’m hard on myself about a lot of things a lot of the time so that I have to kind of work through it and make sure I stay in a positive place”, “I tend to be too critical with myself so sometimes I think I am not good at all. So I agree”.</td>
</tr>
<tr>
<td><strong>Self-regulation</strong> Restoring</td>
<td>“And you go up again and grow up and achieve probably learn and everything and get better”; “then I get a book and learn more about job search (self-enhancement); “maybe today I can recover my status”</td>
</tr>
<tr>
<td><strong>Normalizing</strong></td>
<td>“this is normal, all humans feel not good at all times, it's human nature, we are not perfect”; “that’s that negative reinforcement that we’re all build with and it’s a matter of controlling that negative voice”; “of course there is some time, some, some depressed time, for everyone, we are all like this so I agree”;</td>
</tr>
<tr>
<td><strong>Compensation</strong></td>
<td>“Strongly disagree because I can see my improvement here and I feel I improved myself a lot. For example, when I first come to Canada I speak little, little English and then I worked part time and went to school to improve my language. And lot things like personality: I used to be very anxious or no tolerance but now I've become better and better, yeah”</td>
</tr>
<tr>
<td>Response processes</td>
<td>Verbatim Responses</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Maintaining</td>
<td>“I felt...not ‘not good’ but ‘not good enough’...at times, well, I think it was not good at all, not sure, maybe not good enough...only at times”</td>
</tr>
<tr>
<td>Retrospective</td>
<td>“... you know, there’s the, on the occasion where I make a decision which seemed perfectly logical at the time, and I look at it later and go: what was I thinking! But then, it’s a matter of controlling that negative voice and being able to say: ok look, yeah, you didn’t know this or you didn’t know that. And I never blame myself if I made the best decision based on the information I had, right? But there were these times, so I have to agree”</td>
</tr>
<tr>
<td>Cognitive processes</td>
<td>Autobiographical memory</td>
</tr>
<tr>
<td></td>
<td>“Little images came up of me like lending my roommates things, or, you know, being a good listener to friends, and, just sort of memories, images pop into mind...” ; “I think ‘no good at all’ because of something happened, and because I know I did something to make people unhappy”</td>
</tr>
<tr>
<td>Self-attributions</td>
<td>“it was the circumstance but under normal circumstances I won’t choose that” ; ” it’s usually a reactive thing so it’s not a, it’s not a constant thread. And I wouldn’t, it’s not, I don’t blame, I mean you blame yourself in the minute so if you get a bit upset with your kids you blame yourself but I don’t go away thinking ‘oh I’m a useless mother, I’m no good at, at my job or I’m a hopeless wife’...it’s not like that”</td>
</tr>
<tr>
<td>Affective processes</td>
<td>Self-related emotionality (shame, guilt)</td>
</tr>
<tr>
<td></td>
<td>“Cognition is not important, I am just immersed in emotion, in feeling bad about myself”, “Because it’s not all the time that I think I’m good at all, you know, sometimes I feel bad because I’m not good at all. I am ashamed. So I go with a 3, with agree”</td>
</tr>
<tr>
<td>Contributing, making a difference</td>
<td>“because I, I see the, the difference I make in people’s lives even when I am not good, you know, whether it’s a job, whether it’s family uh and elsewhere like I do a lot of volunteer work as well and so, you know, when you help and do those kinds of things you see the difference that it makes to other people, right?”, “I still make a difference in others ‘lives even if I feel not good at times...so that is good, I can still contribute”</td>
</tr>
</tbody>
</table>
**Item 3: I feel that I have a number of good qualities.**

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-evaluation</td>
<td>Temporal comparisons</td>
</tr>
<tr>
<td></td>
<td>“As I’ve got older I’ve got better so I think: ok so I’m forty nine now so maybe by the time I’m fifty nine I’ll be even better than that”</td>
</tr>
<tr>
<td>Self-improvement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“I think that there are things that I could do better. I think that, that my good qualities are probably the same, a lot of the same things that other people have, I don’t think that I’m exceptional at a lot of things but I think the good qualities that I have are things that people would value, other people would value and that I’m, you know, friendly and I work hard and I try to help people, and that I try to see the good in other people, but I’m not, I could be better so that’s why I wouldn’t give myself ‘I strongly agree.’; Disagree because, you know, you always want to, you always want to improve yourself, right?”</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>Normalizing</td>
</tr>
<tr>
<td></td>
<td>“All humans have some good qualities, hence I must have some good qualities too”; ;” I feel I have a number of good qualities. Yep. Agree because this description is kind of moderate; so everyone has some good qualities”</td>
</tr>
<tr>
<td></td>
<td>“I get that feedback from family and friends and at work, you know, in terms of the kind of character that I’ve build. So that’s why I’m, I feel like I have a number of good qualities because that, I get that feedback from other people. It’s not my ego only”; “So you look for that feedback and if you feel like people are saying: well, you know, you could improve in this area, improve on that area. I mean definitely I know there are areas I can improve in but, you know, that’s based on, you know, on your qualities that you can build on, right”, “I strongly agree, only because I know – people tell me, and I feel I’m a pretty...pretty good person.”</td>
</tr>
<tr>
<td>Socially desirable</td>
<td>desirably responding</td>
</tr>
<tr>
<td></td>
<td>“of course I have weakness I don’t tell you now”</td>
</tr>
<tr>
<td>Cognitive processes</td>
<td>Counterfactual thinking</td>
</tr>
<tr>
<td></td>
<td>“If I did not have any good qualities, I would not be here”</td>
</tr>
<tr>
<td>Response processes</td>
<td>Verbatim Responses</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Declarative</td>
<td>“I am a people person”; “I am kind and good and helpful”, “I am an okay person”</td>
</tr>
<tr>
<td>self-knowledge</td>
<td></td>
</tr>
<tr>
<td>Contributing,</td>
<td>“My life has been taking care, and nurturing other people, and still getting to where I need to go. And so for me that’s, that’s probably the best attribute: I always watch out for others”, “Helping to build people, this is my top good quality”, “others gravitate towards me”</td>
</tr>
<tr>
<td>making a difference</td>
<td></td>
</tr>
</tbody>
</table>
**Item 4: I am able to do things as well as most other people.**

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong></td>
<td>“I’m able to do a lot of the same things without that formal education or that level of education that my coworkers have”; “Actually for the most part I am able to do them better than others who should be more competent than me”.</td>
</tr>
<tr>
<td>Upward social comparisons</td>
<td></td>
</tr>
<tr>
<td><strong>Downward social comparisons</strong></td>
<td>“I am able to do things better than most people”, “I’m so blessed I have all these things that other people wouldn’t have”</td>
</tr>
<tr>
<td><strong>Self-regulation</strong></td>
<td>“My strong agreement with that statement is because I’m getting reinforced from people around me”; “People within my workplace certainly respect me for what I’m doing and what I know and they would come to me so I must be as able as them or more able because they’re coming to me to consult with me”.</td>
</tr>
<tr>
<td>Looking for others’ feedback</td>
<td></td>
</tr>
<tr>
<td><strong>Balancing</strong></td>
<td>“When I meet more and more people I found so usually when I evaluate people I see both sides, both positive side and the negative side of the people. I see the stress of the people and so I think for certain things some people did really very, very good so there’s always people can do better than me but I can do some things not that bad so I don’t agree nor disagree. I see both sides”.</td>
</tr>
<tr>
<td><strong>Quantitative reasoning</strong></td>
<td>“I am an average person”, “I’m a normal person in the middle”, “On average, most of the time I can do things as well as everyone else but then there are situations when I cannot”</td>
</tr>
<tr>
<td><strong>Affective processes</strong></td>
<td></td>
</tr>
<tr>
<td>Gratitude</td>
<td>“I did it. I am proud of it”</td>
</tr>
<tr>
<td>Feeling proud for being able to accomplish things</td>
<td>“I feel incredibly blessed...just grateful”</td>
</tr>
</tbody>
</table>
**Item 5: I feel I do not have much to be proud of.**

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong></td>
<td></td>
</tr>
<tr>
<td>Temporal comparisons</td>
<td>“I have a great deal to be proud of, ’cause I have brought myself back from where I was, so I am very proud of my achievements there”</td>
</tr>
<tr>
<td><strong>Self-expectations</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“I have some higher expectations so um, so even when I reach some things are good I’m not that proud of because I already expect I should be able to do that. But I know uh it’s not as easy as I thought”</td>
</tr>
<tr>
<td><strong>Self-regulation</strong></td>
<td></td>
</tr>
<tr>
<td>Recalibrating the response scale</td>
<td>“Can I choose in between 2 and 3? I do have things to be proud of so it is not that I don’t have much, I don’t want to say that, but I still have a lot to do, to accomplish. So I am really between disagree and agree”</td>
</tr>
<tr>
<td><strong>Quantitative reasoning</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“... my family, my job, you know, all the work I do with people around me so I feel like I’ve got lots to be proud of”; that’s why I gave it a strongly disagree because I feel I have a lot of things to be proud of rather than, that things that I’m unhappy about”</td>
</tr>
<tr>
<td><strong>Self-attribution</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“I am a good person”; “Trying to be a good person and to do things for people make you feel proud”; “I always try to, to be uh a best guy, to make the best decisions”</td>
</tr>
<tr>
<td><strong>Affective processes</strong></td>
<td></td>
</tr>
<tr>
<td>Feeling proud for being able to accomplish things</td>
<td>“I did it. I am proud of it”</td>
</tr>
<tr>
<td><strong>Conflict emotions-thoughts</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“I know I should be proud of but I just sometimes lack of that emotion”; “… there are two sides. As a person, as on emotional side I have things that I should feel proud of but on a different side like being more practical, I have, I’m learning all of those things. So I guess I’m going to write: kind of I don’t know. In between. Yeah, in between. But there’s no in between here (on the RSES)”</td>
</tr>
<tr>
<td><strong>Cognitive processes</strong></td>
<td></td>
</tr>
<tr>
<td>Autobiographical memory</td>
<td>“I have done so many things and I’m proud about that. Both achievements and also taking care about other people like for instance when I was around, I was like twelve I took care of my niece for a long time”</td>
</tr>
<tr>
<td>Response processes</td>
<td>Verbatim Responses</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>Self-improvement</td>
</tr>
<tr>
<td></td>
<td>“I learned from it so next time I will know and will do better”</td>
</tr>
<tr>
<td>Normalizing via</td>
<td></td>
</tr>
<tr>
<td>quantitative reasoning</td>
<td>“There are some times, this is normal, so I agree”</td>
</tr>
<tr>
<td>Keeping oneself</td>
<td></td>
</tr>
<tr>
<td>busy</td>
<td>“I always feel I’m busy so I don’t have any time to say oh I’m useless. Even if I’ve failed at something</td>
</tr>
<tr>
<td></td>
<td>or haven’t done well there’s always lots of other things that I’m doing ok at or even better”</td>
</tr>
<tr>
<td>Looking for</td>
<td></td>
</tr>
<tr>
<td>feedback from others</td>
<td>“Strongly disagree, again for the same reasons because I see all the feedback that I get from people</td>
</tr>
<tr>
<td></td>
<td>where they say that I do help them make a difference”</td>
</tr>
<tr>
<td>Perspective taking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Overall, if I count there are more instances when I did not feel useless”; “Sometimes when I think</td>
</tr>
<tr>
<td></td>
<td>about the world in its sort of enormity and history and our place in the universe, I guess so, yeah,</td>
</tr>
<tr>
<td></td>
<td>we are pretty useless”</td>
</tr>
<tr>
<td>Contrasting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“in contrast with everything I said- high self-esteem-there were moments when I felt helpless”</td>
</tr>
<tr>
<td>Self-attribution</td>
<td>Temporary</td>
</tr>
<tr>
<td>Situational</td>
<td>“I feel temporarily inadequate so I don’t feel useless but I wouldn’t just sit there and think: oh</td>
</tr>
<tr>
<td></td>
<td>I can’t help with this”</td>
</tr>
<tr>
<td>Anticipatory</td>
<td>Intentional</td>
</tr>
<tr>
<td>Processes</td>
<td>“Even if something bad happens I know I’ll be depressed for some time but I know I, I have that thing,</td>
</tr>
<tr>
<td></td>
<td>that I’ll again get up from that. Maybe it will take more time or lesser time but I know I will”;</td>
</tr>
<tr>
<td></td>
<td>“There is nothing that can shatter me emotionally. The only thing that, like many times disturbed me</td>
</tr>
<tr>
<td></td>
<td>is like when I feel that people are not considerate, that’s the main thing”</td>
</tr>
</tbody>
</table>
**Item 7:** I feel that I’m a person of worth, at least on an equal plane with others.

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong></td>
<td>“I am a fair person”, “I am a good, kind person”; “I uphold myself well. Yeah I don’t put myself down like to that level. But I’m kinda like an average...yeah, like I have like a good amount of self-esteem and stuff”.</td>
</tr>
<tr>
<td><strong>Abstract, philosophical-religious beliefs</strong></td>
<td>“I don’t need to contribute to be a person of worth; worth is by birth”, “Being alive is valuable in itself; I am a person of worth just by being alive”</td>
</tr>
<tr>
<td><strong>Implicit theories of a good person</strong></td>
<td>“Everyone of us is a unique blend. Um, you know, so I look at it and say there, there’s an old saying about if everybody’s the same it would be a boring world, right? So I, I look for, we’re all different pieces of the puzzle [Mhm] how do we fit together to make things work?”</td>
</tr>
<tr>
<td><strong>Social comparisons</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Self-regulation</strong></td>
<td>“There are two things: what I feel and what actually is. I feel I am totally worth but in practical then I go or I over criticize myself and I think it’s not I over criticize myself. I think it’s mainly because I am low in emotional confidence inside me [I need more of like pampering and loving, from the inside. Like self-compassion or self-love”</td>
</tr>
<tr>
<td><strong>Quantitative reasoning</strong></td>
<td>“I agree because it says ‘at least on an equal basis with others.’ You know so I think I’m pretty much in the average bunch there”.</td>
</tr>
<tr>
<td><strong>Self-attribution</strong></td>
<td>“I am a good person”; “Trying to be a good person and to do things for people make you feel proud”; “I always try to, to be uh a best guy, to make the best decisions”</td>
</tr>
</tbody>
</table>
**Item 8: I wish I could have more respect for myself.**

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong></td>
<td><strong>Temporal comparisons</strong></td>
</tr>
<tr>
<td><strong>Social comparisons</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Self-regulation</strong></td>
<td><strong>Looking for others’ feedback or reassurance</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Recognizing the effort not just the outcome</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Self-improvement</strong></td>
</tr>
<tr>
<td><strong>Affective processes</strong></td>
<td><strong>Anger</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Regret</strong></td>
</tr>
<tr>
<td><strong>Cognitive processes</strong></td>
<td><strong>Autobiographical memory</strong></td>
</tr>
<tr>
<td><strong>Moral agentic evaluation</strong></td>
<td></td>
</tr>
</tbody>
</table>
### Item 9: All in all, I am inclined to feel that I am a failure.

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-evaluation</td>
<td>Downward comparisons</td>
</tr>
<tr>
<td></td>
<td>“People who live a hand to mouth existence, addicts, someone who has nobody because they don’t know how to keep relationships, that is a failure. In comparison with these people I am not a failure, strongly disagree”</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>Harmonizing with others</td>
</tr>
<tr>
<td></td>
<td>“Failure is if you cannot get along, ummm, harmony, harmonize with others. Living just by yourself. Selfish, you know”</td>
</tr>
<tr>
<td></td>
<td>Comparing and contrasting achievements and personal integrity</td>
</tr>
<tr>
<td></td>
<td>“Living in the big house, driving the fancy car, all those little status symbols they don’t mean success. You can have all of those and still be a failure”</td>
</tr>
<tr>
<td>Cognitive processes</td>
<td>Autobiographical memory</td>
</tr>
<tr>
<td></td>
<td>“I’m afraid of I couldn’t do that. Yeah afraid of failure. Yeah, that, you know, when at school take exams and the university”; “Sometimes I lack confidence, I feel I could not do it, yeah, I’m afraid that I will fail”</td>
</tr>
<tr>
<td>Moral agentic evaluation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“…and it’s not what they have accomplished necessarily that matters, there’s an awful lot of people who have made it to quote ‘the corner office’ and, as a friend of mine introduced me to a saying, he says: they’re a waste of skin. You know, as human being goes, they have very little in the way of redeeming qualities. No moral compass, no values”</td>
</tr>
<tr>
<td>Contributing self</td>
<td>“People who made the world a better place by being here; that’s a true success in life, because they can bring this joy and that energy to, to others, throughout the world”, “I feel like I’ve done so much, like not for myself, but like for others and for myself”</td>
</tr>
</tbody>
</table>
**Item 10: I take a positive attitude toward myself.**

<table>
<thead>
<tr>
<th>Response processes</th>
<th>Verbatim Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-evaluation</strong></td>
<td>“People who live a hand to mouth existence, addicts, someone who has nobody because they don’t know how to keep relationships, that is a failure. In comparison with these people I am not a failure, strongly disagree”</td>
</tr>
<tr>
<td>Downward comparisons</td>
<td></td>
</tr>
<tr>
<td><strong>Self-regulation</strong></td>
<td>“Failure is if you cannot get along, ummm, harmony, harmonize with others. Living just by yourself. Selfish, you know”</td>
</tr>
<tr>
<td>Harmonizing with others</td>
<td></td>
</tr>
<tr>
<td><strong>Comparing and contrasting achievements and personal integrity</strong></td>
<td>“Living in the big house, driving the fancy car, all those little status symbols they don’t mean success. You can have all of those and still be a failure”</td>
</tr>
<tr>
<td><strong>Cognitive processes</strong></td>
<td>“I’m afraid of I couldn’t do that. Yeah afraid of failure. Yeah, that, you know, when at school take exams and the university”; “Sometimes I lack confidence, I feel I could not do it, yeah, I’m afraid that I will fail”</td>
</tr>
<tr>
<td>Autobiographical memory</td>
<td></td>
</tr>
<tr>
<td><strong>Moral agentic evaluation</strong></td>
<td>“…and it’s not what they have accomplished necessarily that matters, there’s an awful lot of people who have made it to quote ‘the corner office’ and, as a friend of mine introduced me to a saying, he says: they’re a waste of skin. You know, as human being goes, they have very little in the way of redeeming qualities. No moral compass, no values”</td>
</tr>
<tr>
<td><strong>Contributing self</strong></td>
<td>“People who made the world a better place by being here; that’s a true success in life, because they can bring this joy and that energy to, to others, throughout the world”; “I feel like I’ve done so much, like not for myself, but like for others and for myself”</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>