

Validating Measures of Self Control via Rasch Measurement

Jonathan Hasford
Department of Marketing, University of Kentucky

Kelly D. Bradley
Department of Educational Policy Studies & Evaluation, University of Kentucky

ABSTRACT

Self control has been offered as a fundamental explanation for many behavioral outcomes. Until recently, measurement of self control had been inadequate, with advances being made only in specific domains. Tangney, Baumeister, and Boone (2004) introduced a reflective measure of self control which has gained popularity in social science research. However, the authors did not subject this critical measure to item response theory, which is a more thorough and detailed analysis in comparison to classical test theory. The following research analyzes their measure of perceived self control with item response theory, and more specifically, Rasch measurement. Findings suggest that the unidimensional construct of self control is captured by the scale but should include revisions for future use.

Self control, or the individually imposed ability to regulate one's behavior, is a leading topic of interest across social science research. Given that theoretical explanations for personal and social problems are often founded in self control issues (Baumeister and Alquist 2009; Baumeister, Heatherton, and Tice 1994), accurate measurement of self-control is a critical issue. Prior attempts to capture self control in social science research are often context specific to a particular behavior (e.g. Brandon, Oescher, and Loftin 1990; Rosenbaum 1980) or only a subscale measure of self control from another concept of interest (e.g. Bearden, Hardesty, and Rose 2001; Gough 1987). Recent research, however, has created a self-report measure of self control which aims to provide an individual difference measure and predict various outcomes in psychological behavior (Tangney, Baumeister, and Boone 2004). Utilizing classical test theory, Tangney et al. (2004) demonstrate their self-report scale of self control items (see Appendix) is correlated with a variety of individual behaviors and psychological well-being, including eating disorders, alcoholism, anxiety, depression, emotional stability, family conflict, and interpersonal relationship quality.

Given the importance of self control in behavioral research and the widespread adoption of the Tangney et al. (2004) measure of self control, the purpose of the following research is to determine whether this measure of perceived self control is acceptable within an item response theory framework, and specifically, Rasch measurement. The Rasch model is a one-parameter logistic model within item response theory (IRT) in which the amount of a given latent trait in a person and the amount of that same latent trait reflected in various items can be estimated independently yet still compared explicitly to one another. Given that the Rasch model follows mathematically from the requirement of invariance of comparisons among persons and items, a Rasch analysis is appropriate when the total score on a questionnaire is used to make inferences

about an individual's level of a latent trait inherent in that individual. Classical Test Theory (CTT) also uses the total score to characterize each person, but it asserts the total score as the relevant statistic with little consideration of anomalies in the items or the respondents.

Applications of the Rasch model account for these anomalies and provide a more informative score (Andrich & Luo, 2003).

The Rasch model yields a comprehensive picture of the construct under measurement and the respondents on that measure. It allows observations of respondents and items to be connected in a way that indicates the occurrence of a response as probability rather than certainty and maintains order in that the probability of providing a response defines an order of respondents and items (Wright and Masters, 1982). It is a static model, meaning that for each person having a certain amount of the given latent trait, it specifies the probability of a response in one of the categories of an item. The Rasch model uses the sum of the item ratings simply as a starting point for estimating probabilities of those responding to each item. In the case of a questionnaire, it is based upon an individual's willingness to endorse a set of items and the difficulty to endorse those items. It is assumed the difficulty to endorse is the main characteristic influencing responses (Linacre, 1999).

Sample

Two separate samples of data were collected which contained a subset of the self control items from Tangney et al. (2004). The first sample contained 66 undergraduate business students in an upper-level marketing course who participated in the study for extra credit. The second sample contained 75 undergraduate students who enrolled in an introductory business course and completed the survey for course credit. No distinguishing characteristics were present between

samples. However, the use of a second sample provides further evidence to the validity and reliability of the Rasch model. Additionally, participants were also combined and analyzed in a stacked sample of all participants to further validate the overall model. If the proposed model of self control is of high quality, measures of fit should be approximately similar across both samples.

Method

Participants answered 20 of Tangney et al.'s (2004) self control items as part of a study on consumption behavior. These items are bolded in the Appendix. Only using a subset of items should not affect the results of analysis, since Tangney et al. (2004) have a longer (36 item) and shorter (13 item) version of the measure, both of which correlate similarly across behavior. The key measures of interest were from responses to each of the 20 items on a scale of Strongly Disagree-1 to Strongly Agree-7. This response scale differs from Tangney et al.'s (2004) five-point scale (Not At All-1 to Very Much-5). The revised response scale is superior in terms of capturing a true interval measure which can be utilized for data analysis.

Results

To determine whether the scale is a quality measure of self control, three main findings from the Rasch analysis are considered. First, item misfit across individual items is considered. Then, the map of person abilities and item difficulties is analyzed. Lastly, person and item reliabilities are presented. Both samples are considered here, with the test of a quality measure involving whether both samples demonstrate minimal item misfit, dispersion across person abilities and item difficulties, and high person and item reliabilities.

Item Misfit. The 20 self control items were analyzed for infit and outfit with the model. Appropriate infit and outfit statistics should lie between .6 and 1.5 for each item (Bond and Fox 2007). Fit statistics which are too high suggest the model is inappropriate for the data. Fit statistics which are too low suggest overfit in the data, meaning that responses are too well-fitting for reality. The results of the item misfit analysis are presented in Table 1. Results demonstrate acceptable fit for almost every item except item 8 in sample A, item 15 in sample B and the combined sample, and item 16 for sample A and B. However, these misfit statistics are only slightly beyond prescribed cutoffs (Bond and Fox 2007) and should not be a primary concern. Therefore, the measures provide good fit within the model and further analysis with all items is appropriate.

Table 1

Sample A			Sample B			Combined	
<u>Infit</u>	<u>Outfit</u>		<u>Infit</u>	<u>Outfit</u>		<u>Infit</u>	<u>Outfit</u>
0.75	0.75	<i>Item 1</i>	0.94	0.92	<i>Item 1</i>	0.84	0.83
1	1.03	<i>Item 2</i>	1.19	1.18	<i>Item 2</i>	1.09	1.1
1.17	1.15	<i>Item 3</i>	1.07	1.09	<i>Item 3</i>	1.16	1.14
1.3	1.35	<i>Item 4</i>	0.89	0.89	<i>Item 4</i>	1.05	1.06
1.09	1.15	<i>Item 5</i>	1.09	1.12	<i>Item 5</i>	1.08	1.12
0.79	0.78	<i>Item 6</i>	0.85	0.88	<i>Item 6</i>	0.86	0.87
0.88	0.94	<i>Item 7</i>	1.05	1.06	<i>Item 7</i>	1.07	1.09
1.53	1.58	<i>Item 8</i>	1.35	1.33	<i>Item 8</i>	1.42	1.41
1	1.08	<i>Item 9</i>	1.08	1.08	<i>Item 9</i>	1.04	1.07
0.99	0.98	<i>Item 10</i>	0.91	0.91	<i>Item 10</i>	0.94	0.94
0.97	0.94	<i>Item 11</i>	1.01	0.98	<i>Item 11</i>	0.97	0.94
0.89	0.89	<i>Item 12</i>	1.03	1.01	<i>Item 12</i>	0.95	0.93
0.76	0.77	<i>Item 13</i>	0.99	1.02	<i>Item 13</i>	0.88	0.91
1.1	1.14	<i>Item 14</i>	0.72	0.72	<i>Item 14</i>	0.89	0.88
1.44	1.46	<i>Item 15</i>	1.53	1.72	<i>Item 15</i>	1.46	1.61
0.56	0.57	<i>Item 16</i>	0.5	0.51	<i>Item 16</i>	0.6	0.6
1.09	1.1	<i>Item 17</i>	0.92	0.93	<i>Item 17</i>	0.98	0.99
0.61	0.61	<i>Item 18</i>	1.26	1.31	<i>Item 18</i>	1.16	1.14
1.07	1.13	<i>Item 19</i>	1	0.99	<i>Item 19</i>	1.01	1.02
0.88	0.87	<i>Item 20</i>	0.92	0.91	<i>Item 20</i>	0.89	0.88

Variable Maps. The variable map provides logit scores across both person abilities and item difficulties which can be compared visually to a normal distribution (Bond and Fox 2007). The variable maps for each study can be seen in Figures 1 through 3. The variable map of the combined sample demonstrates similar results to the sample B variable map, so further discussion of the findings will only compare sample A with sample B. Analysis of the variable maps demonstrates sufficient distribution across person abilities and item difficulties. Both samples follow a pattern of normal distribution in which most individuals have moderate levels of self control and few have high or low levels. However, the positioning of items between samples raises some issues. First, item 17 in sample A, which distinguishes among individuals with higher amounts of self control, is actually the lowest measure of self control in sample B. Furthermore, some items (such as item 1 and item 15) predict higher levels of self control in one sample and lower levels of self control in the other. Lastly, while sample A has an approximate match between the distribution of people across items, sample B does not have enough items that distinguish among individuals with very high levels of self control. These issues question the set of items used to formulate a measure of self control. A reanalysis of the self control with the full scale may be in order to determine whether better fit between persons and items across the variable map is present. Inclusion of the remaining items may increase the ability to differentiate perceived self control at high levels within individuals. Further consideration of these issues is provided in the discussion.

Figure 1 – Sample A Variable Map

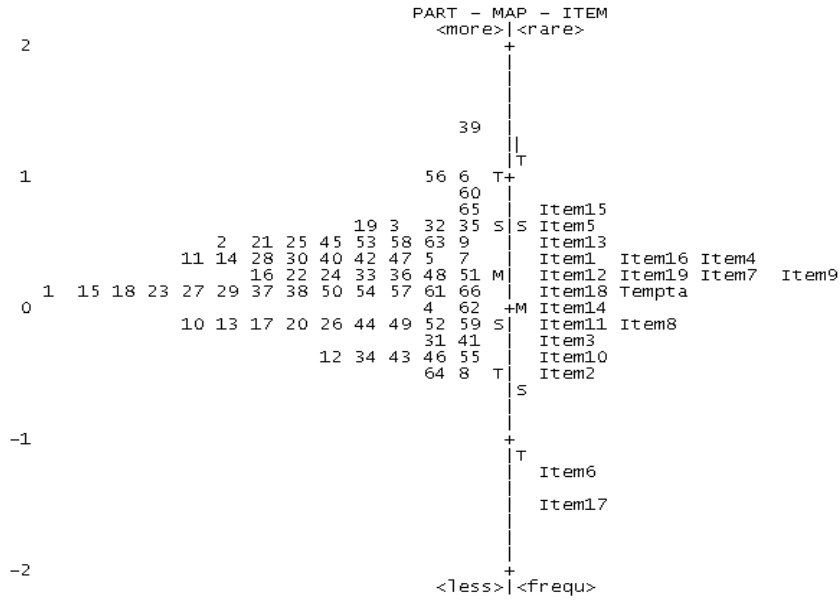
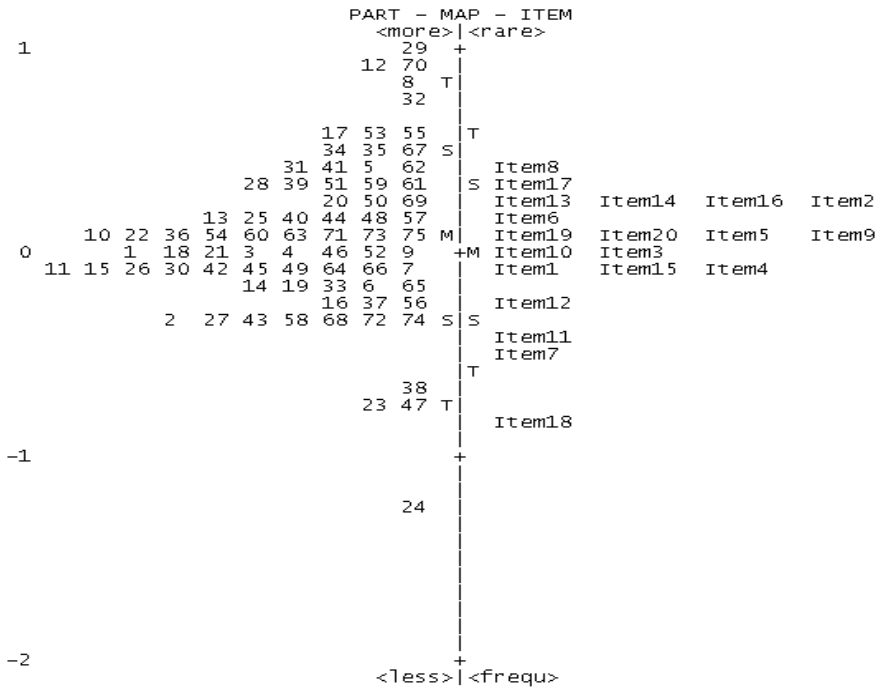


Figure 2 – Sample B Variable Map



Discussion

Self control is a topic of interest in the social sciences, given its ability to impact a number of choices and behaviors. In this research, a leading measure of perceived self control in the psychological literature is investigated. Results demonstrate that, across the available items in the study, the self control measure developed by Tangney et al. (2004) appears to be adequate. All 20 items are either within acceptable misfit measures or close enough to retain inclusion in the scale. Furthermore, person and item reliabilities across samples are high and nearly identical to one another, suggesting the scale's ability to produce consistent results. However, widespread implementation of the self control scale should occur with some trepidation.

The major concern with the self control scale which emerged from analysis of the variable maps is that the scale does not adequately measure very high or very low levels of self control. In the variable maps for the samples, the items were distributed within +/-1.5 logits of the midpoint of ability. This suggests that individuals who are highly (not) capable of exercising self control are not being captured by this measure. In the 2nd and combined samples, the variable maps clearly demonstrated that more measures of high self control were necessary to distinguish among individuals with high amounts of self control (i.e. the top portions of the map). However, this could be an issue stemming from only having a slight majority of the items from the original scale for analysis. Subsequent research could determine whether this issue is resolved by including all items in the scale.

Furthermore, the positioning of certain items between samples raises some concerns. As previously discussed, certain items were related to high self control in one sample and low self control in the other. This raises questions of reliability at the item level, which is a major factor that Rasch analysis aims to demonstrate (Bond and Fox 2007). Once again, however, this could

possibly be a function of the limited number of items from the original scale. Having all items may shift the 20 included here either upward or downward on the variable map, which would restore consistency in the measure. Another possible explanation may relate to the nature of the scale. Since it aims to demonstrate general self control across a variety of behaviors, fluctuation may be expected within items which still capture the overall aim of the measure. Further research should investigate whether dimensions of self control exist and how they combine to form an overall measure of self control.

The self control scale developed by Tangney et al. (2004) appears to be an acceptable measure of self control. However, analysis of the full and short-form self control scales may provide additional insights not found here. For future studies involving individual differences in self control, this measure should be considered adequate. Moving forward, the topic of self control in general should remain at the forefront of social science research given its ability to impact a wide variety of general life behaviors and outcomes.

References

- Andrich, D. & Luo, G. (2003). Conditional Pairwise Estimation in the Rasch Model for Ordered Response Categories using Principal Components. *Journal of Applied Measurement*, 4 (3), 205-221.
- Baumeister, R. F. & Alquist, J. L. (2009). Is There a Downside to Good Self-control? *Self and Identity*, 8 (2), 115-130.
- Baumeister, R. F., T. F. Heatherton, and D. M. Tice (1994), *Losing Control: How and Why People Fail at Self-Regulation*. San Diego: Academic Press, Inc.
- Bearden, W. O., D. M. Hardesty, and R. L. Rose (2001). Consumer Self-Confidence: Refinements in Conceptualization and Measurement. *Journal of Consumer Research*, 28 (1), 121-134.
- Bond, T. G. & Fox, C. M. (2007), *Applying the Rasch Model: Fundamental Measurement in the Human Sciences (2nd Edition)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Brandon, J. E., J. Oescher, and J. M. Loftin (1990). The Self-Control Questionnaire: An Assessment. *Health Values*, 14, 3-9.
- Gough, H. G. (1987), *California Psychological Inventory – Administrator’s Guide*. Palo Alto, CA: Consulting Psychologists Press.
- Linacre, J. M. (1999). *A User’s Guide to Facets Rasch Measurement Computer Program*. Chicago, IL: MESA Press.
- Rosenbaum, M. (1980). A Schedule for Assessing Self-Control Behaviors: Preliminary Findings. *Behavior Therapy*, 11, 109-121.
- Smith, E. V. (2004). Evidence for the Reliability of Measures and Validity of Measure Interpretation: A Rasch Measurement Perspective. E. V. Smith and R. M. Smith (eds.), *Introduction to Rasch Measurement*. Maple Grove, MN: JAM Press.
- Tangney, J. P., R. F. Baumeister, and A. L. Boone (2004). High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success. *Journal of Personality*, 72 (2), 271-324.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.

Appendix

The Self Control Scale developed by Tangney et al. (2004). Items included in Rasch analysis are bold. Items with an * represent the short form of the self control scale. Items with an (R) are reverse coded.

- 1. I am good at resisting temptation ***
- 2. I have a hard time breaking bad habits (R) ***
- 3. I am lazy (R) ***
- 4. I say inappropriate things (R) ***
- 5. I never allow myself to lose control**
- 6. I do certain things that are bad for me, if they are fun (R) ***
- 7. People can count on me to keep on schedule**
- 8. Getting up in the morning is hard for me (R)**
- 9. I have trouble saying no (R)**
- 10. I change my mind fairly often (R)**
- 11. I blurt out whatever is on my mind (R)**
- 12. People would describe me as impulsive (R)**
- 13. I refuse things that are bad for me**
- 14. I spend too much money (R)**
- 15. I keep everything neat**
- 16. I am self-indulgent at times (R)**
- 17. I wish I had more self-discipline (R) ***
- 18. I am reliable**
- 19. I get carried away by my feelings (R)**
- 20. I do many things on the spur of the moment (R)**
21. I don't keep secrets very well (R)
22. People would say that I have iron self- discipline *
23. I have worked or studied all night at the last minute (R)
24. I'm not easily discouraged
25. I'd be better off if I stopped to think before acting (R)
26. I engage in healthy practices
27. I eat healthy foods
28. Pleasure and fun sometimes keep me from getting work done (R) *
29. I have trouble concentrating (R) *
30. I am able to work effectively toward long-term goals *
31. Sometimes I can't stop myself from doing something, even if I know it is wrong (R) *
32. I often act without thinking through all the alternatives (R) *
33. I lose my temper too easily (R)
34. I often interrupt people (R)
35. I sometimes drink or use drugs to excess (R)
36. I am always on time.