10-19-2012

Using Rasch Residuals to Analyze the Demographic Characteristics of Respondents with Unexpected Incorrect Answers: Implications for Construct Validity

Robert K. Gable
Johnson & Wales University - Providence, rgable@jwu.edu

Larry Ludlow, Ph.D.
Boston College, ludlow@bc.edu

Stacey L. Kite
Johnson & Wales University, skite@jwu.edu

Using Rasch Residuals to Analyze the Demographic Characteristics of Respondents with Unexpected Incorrect Answers: Implications for Construct Validity

Robert K. Gable
Center for Research and Evaluation
Alan Shawn Feinstein Graduate School
Johnson & Wales University

Larry H. Ludlow
Educational Research
Lynch School of Education
Boston College

Stacey L. Kite
Center for Research and Evaluation
Alan Shawn Feinstein Graduate School
Johnson & Wales University

---

[1]Paper presented at the 43rd annual meeting of the Northeastern Educational Research Association, October 19, 2012, Rocky Hill, CT.

Abstract

This paper illustrates the use of Rasch model residuals to better understand perceived respondent meaning and structure of item content in the context of construct validity. Data were gathered from1366 grade 7-8 students using the *Survey of Knowledge of Internet Risk and Internet Behavior*. The characteristics of the respondents with unexpected patterns of residuals for persons not fitting the Rasch model are examined for 7 items defining the Knowledge of Internet Risk scale. These analyses contribute to a better understanding of the item content and person scores, and contribute to more meaningful score inferences.

Purpose/Framework

The purpose of this paper is to illustrate the use of Rasch model residuals to better understand item content and person characteristics. As described in Gable, Ludlow, and Wolf (1990) and Gable and Wolf (1993), the perceived respondent meaning and structure of item content and the characteristics of the respondents can be studied by examining unexpected patterns of residuals for persons not fitting the Rasch model. These analyses can contribute to a better understanding of the item content and person scores, and contribute to more meaningful inferences from test scores in the context of construct validity.

Methodology

Instrumentation and Data Analysis

The instrument used to gather the *N*=1366 grade 7-8 data was the *Survey of Knowledge of Internet Risk and Internet Behavior*. The following sections from Gable, Ludlow, McCoach, and Kite ((2011) describe the 26 item survey, noting that this paper will only focus on the Rasch residual analyses for the 7 Knowledge of Interent Risk items.

> *Scales and scoring technique*. The survey contained 7 literature derived demographic items: gender, grade level, have older siblings, earn good grades, are you popular, ever get into trouble at school, and own a cell phone (e.g., Shariff, 2008) and 26 statements constructed to describe students' knowledge of risks and behaviors associated with using the Internet (Gable, Ludlow, Kite, McCoach, & Filippelli, 2009).

Response format. Students were asked to *Agree* or *Disagree* with each statement. Responses were scored "1" or "0" to reflect a high level of the attribute measured by the scale. Appropriate agreeing or disagreeing with a statement received a score of "1" (e.g., agree with the statement: *Making threats online can get me into trouble with the police.*); an inappropriate *agree* or *disagree* response was scored "0". This scoring technique was designed to produce scores where high scoring students had higher levels of knowledge of Internet risks.

The Internet "Knowledge" scale was composed of 7 items describing knowledge of appropriate behavior on social networks and potential risk of Internet predators. The Knowledge items were intended to span a unidimensional, hierarchically ordered continuum consistent with the Rasch measurement model.

Validity. Content validity of the items was supported through the cyberbullying literature (e.g., Hinduja & Patchin, 2009) and judgmental review by *N*=5 middle school teachers and *N*=2 principals. Construct validity was examined through Rasch model.

Rasch model. The Rasch model analysis served as a confirmatory test of the extent to which the knowledge items successfully defined a unidimensional, hierarchically ordered scale of Internet Knowledge of Internet risks. The objective of the analysis reported here was to examine item and person goodness-of-fit statistics, i.e., the extent to which the observed responses were expected under the model.

Reliability. Cronbach's alpha for the Knowledge scores was .69. The use of the binary (*Agree, Disagree*) response format most likely contributed to the lower than desired reliability level because of the resulting restriction on item and scale variance. (p. 219 - 221).

Results and Discussion

Evidence of Internal Structure: Construct Validity

Support for the internal structure or construct validity of the Knowledge scale score interpretations for these data was reported by Gable, Ludlow, McCoach, Kite, & Filippeli (2009) using Rasch model analyses. Figure 1 presents the "Wright" variable map for the Knowledge items. Support for construct validity is present as the items clearly span the "difficulty" continuum, which allows for finer interpretations of high and low scoring students. Inspection of the variable map indicates that item v2b (*Making threats online can get me into trouble with the police.*) is the easiest item to agree with and item v39b (*An internet predator could contact me based on what my friends have posted about me*) is the hardest or most difficult item to agree with.

Table 1 contains the scoring for the 7 Knowledge items. Three of the items (v22b, v29b, and v39b) were associated with low levels of knowledge of the risks regarding contact by an Internet predator using information posted online by the student or the students' friends. For these three items less than 30% of the students offered the appropriate response.

Follow-up Analysis of Rasch Residuals

The difference between an observed and expected response can be expressed as a standardized residual (Wright & Stone, 1979).The residuals were transformed into standardized weighted fit statistics and their unstandardized mean square version. These two summary statistics are generally sufficient to reveal consistent unexpected responses for both individual items and respondents (Gable, Ludlow, McCoach, Kite, & Filippeli, 2009).

Table 2 contains the item statistics and misfit order for the 7 Knowledge items. Note that item v2b (*Making threats online can get me into trouble with the police*.) located on the far right side of the first row in the table. This item displayed moderate misfit (Outfit MNSQ=1.86) due to large numbers of students giving a surprising *disagree* incorrect response to a relatively easy item to agree with. The item misfit for item v2b was examined using the most misfitting response strings listed in Table 3. In this table the first row lists the most misfitting item (v2b), the item misfit statistic (1.86), and then the actual responses (0=incorrect answer) of the most misfitting students. The person numbers listed in a vertical manner with the 0 entries below them contain the *n*=6 students of interest (i.e., students 1354, 1298, 1147, 820, 481, 354, and 328). These *n*=6 misfitting students were selected by examining the "student misfit order statistics" presented in Table 4. Note that the "Entry Number" or student ID listed in ascending order in the first 6 rows of Table 4 identifies the *n*=6 target students. Their listed "raw score" and "count" indicate that they were high Knowledge students with 6 of 7 correct answers; we know from Table 3 that they had incorrect answers to item v2b. These were students with a high level of knowledge of appropriate behavior (i.e., correct/agree) answers for the remaining 6 items defining the scale), but gave an unexpected incorrect answer (disagree) to this item.

Using the process described by Gable, Ludlow, McCoach, Kite, & Filippeli (2009), follow-up examination of the content of item v2b and the demographic characteristics of the students were examined. We felt that the item content did not provoke the unexpected response. Therefore, the following demographic characteristics of the students were examined: school type, gender, grade level, having an older sibling, getting good grades, perception of popularity with friends, getting into trouble, and owning a cell phone. These demographics were selected for the survey based on the cyberbully literature.

Descriptive statistics were run for each demographic variable for the *n*=6 students and the remaining *n*=1360 respondents. While the data in Table 5 are only descriptive in nature, examination of the demographic characteristics suggested that, when compared to the remaining *n*=1360 students, the students with unexpected incorrect responses had the following characteristics (see boxed percents): tended to be males in grade 8, have an older sibling, feel they are popular, indicated they get in trouble, and own a cell phone.

Educational Implications

The researchers are discussing the implications of these findings as they contribute to the interpretations of high and low scoring students, the central issue of construct validity. Could it be that giving an incorrect *disagree* answer to the "easiest" item defining the scale suggests an attempt to hide their possible activities on the Internet? Social desirability or a feeling of threat, due to the mention of police may have been a factor.

The item was deemed to be in-line with both the literature and the other items on the instrument. While these findings are interesting in a descriptive manner (see below), they do not convey any meaningful and consistent response patterns that would suggest that this particular item is flawed. Thus, we concluded that there was not sufficient construct validity evidence to eliminate item v2b from the Knowledge scale. And, in fact, leaving v2b in could unobtrusively provoke exactly the kind of "at-risk" response linked to personal undesirable behavior we hope to uncover and then be able to offer an intervention for.

References

Gable, R. K., Ludlow, L. L., Kite, S. L., McCoach, D. B., & Filippelli, L. P. (April, 2009). *Development and validation of the Survey of Knowledge of Internet Risk and Behavior.* Paper presented at the annual meeting of the American Educational Research Association, San Diego.

Gable, R. K., Ludlow, L. H., McCoach, D. B., & Kite, S. L. (2011). Validation of the Survey of Knowledge of Internet Risk and Internet Behavior. *Educational and Psychological Measurement, 71*(1), 217-230.

Gable, R. K., Ludlow, L. H., & Wolf, M. D. (1990). The use of classical and Rasch latent trait models to enhance the validity of affective scales. *Educational and Psychological Measurement*, *50*(3), 869-878.

Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain.* Boston: Kluwer Academic Publishers. (now Springer)

Hinduja, S., & Patchin, J. W. (2009). Bullying beyond the schoolyard: Preventing and responding to cyberbullying. Thousand Oaks, CA: Corwin Press.

Shariff, S. (2008). *Cyberbullying: Issues and solutions for the school, the classroom, and the home*. New York: Routledge.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.

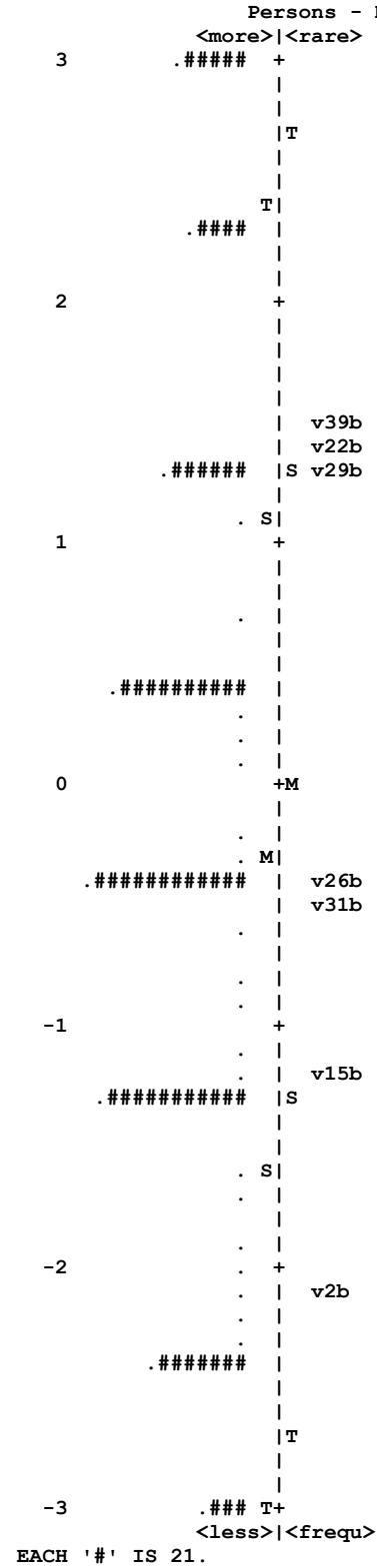

Figure 1. Knowledge Scale: Map Showing People and Item Locations (Gable, Ludlow, Kite, & Filippelli, 2009)

Table 1

Percent of Agree and Disagree Responses for Items Defining the Final Scales

Scale/Item	Percent[1] Agree	Disagree
Knowledge		
Making threats online can get me in trouble with the police. (v2b)	[77]	23
An on-line predator could contact me using a social networking site like myspace or facebook if I posted my personal information on it. (v15b)	[65]	35
With the contact information I put on myspace or facebook, it would be easy for an internet predator to contact me. (v22b)	[27]	73
An internet predator can easily use sites such as Google earth, MSN live or other programs to locate my school and house. (v26b)	[52]	48
An internet predator could make contact with me based on the information I have posted online. (v29b)	[29]	71
Threats online that I carry out at school can get me into trouble. (v31b)	[55]	45
An internet predator could contact me based on what my friends have posted about me. (v39b)	[26]	74

[1]Appropriate or "correct" answers are boxed. Scoring key: correct answers are scored 1 and incorrect answers 0.

Table 2

Knowledge Scale Item Statistics and Misfit Order

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
1	932	1170	-2.11	.08	1.15	3.3	1.86	6.4	A .45	78.8	81.4	v2b
2	768	1166	-1.20	.07	1.03	.9	1.20	2.8	B .54	73.3	74.2	v15b
6	624	1160	-.52	.07	1.03	.9	1.11	2.2	C .57	72.5	72.3	v31b
4	595	1159	-.39	.07	.97	-.9	1.01	.2	D .60	72.8	71.5	v26b
3	251	1162	1.44	.08	.99	-.3	.94	-.6	c .62	81.4	82.1	v22b
7	240	1162	1.52	.08	.90	-2.1	.91	-.9	b .65	83.9	82.7	v39b
5	277	1163	1.27	.08	.79	-5.3	.68	-4.1	a .69	86.8	80.5	v29b
MEAN	526.7	1163.1	.00	.08	.98	-.5	1.10	.9		78.5	77.8	
S.D.	255.6	3.5	1.33	.01	.10	2.5	.34	3.1		5.4	4.6	

Table 3

Most Misfitting Student Response Strings

```
Item         OUTMNSQ  |Person
                      |11111                  111
                      |3210099888876643321  321876655433322222   6843311
                      |5946390432123085289  59315869427119865186398780949
                      |48755363702974148725139781686715738180493383377416
                    high-------------------------------------------------
     1 v2b      1.86 1|0.0......0....000....00000000000000000000........
     2 v15b     1.20 2|...000..0.........0000.....0000.............  .......
     6 v31b     1.11 6|.0....00..0000.......0.........0..00..............
     4 v26b     1.01 4|.....................0.............00.0..........
     3 v22b      .94 3|.............................................1...11
     7 v39b      .91 7|.............................................1..111..
     5 v29b      .68 5|.............................................1......
                      |-----------------------------------------------low-
                      |11111998888766433215111187665543332222286368433119
                      |3210090432123085289  32115869427119865193398780946
                      |5946336370297414872  593816867157381804   8337741
                      |48755                397
```

Table 4

Student Misfit Order Statistics

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	Person
328	6	7	2.34	1.15	1.63	1.0	9.90	3.3	A-.65	85.7	85.7	328118
354	6	7	2.34	1.15	1.63	1.0	9.90	3.3	B-.65	85.7	85.7	354118
481	6	7	2.34	1.15	1.63	1.0	9.90	3.3	C-.65	85.7	85.7	481228
820	6	7	2.34	1.15	1.63	1.0	9.90	3.3	D-.65	85.7	85.7	820216
1147	6	7	2.34	1.15	1.63	1.0	9.90	3.3	E-.65	85.7	85.7	1147316
1354	6	7	2.34	1.15	1.63	1.0	9.90	3.3	F-.65	85.7	85.7	1354317

Table 5

Demographic Profile of Higher Knowledge Students Who Responded with an Unexpected Incorrect Answer for Item 2:

Making threats online can get me in trouble with the police.[1]

Demographics Variables	Target Group (*n* =6)		Remaining Students (*n* =1360)	
	F	%	F	%
School				
Urban	2	33	478	35
Suburban	2	33	416	31
Rural	2	33	466	34
Gender				
Male	5	83	693	51
Female	1	17	665	49
Grade Level				
6	2	33	364	27
7	1	17	497	37
8	3	50	498	37
Older Sibling				
Yes	4	67	896	66
No	2	33	460	34
Good Grades				
Yes	3	50	1120	84
No	3	50	214	16
Popular with Friends				
Yes	5	83	1112	83
No	1	17	234	17
Get in Trouble				
Yes	4	67	637	47
No	2	33	711	53
Own a Cell Phone				
Yes	5	83	962	71
No	1	17	395	29

[1]These *n* =6 students had correct answers for 6 of 7 Knowledge items and an unexpected incorrect answer to item v2b.