Examining the Association between Patient-Reported Symptoms of Attention and Memory Dysfunction with Objective Cognitive Performance: A Latent Regression Rasch Model Approach

Yuelin Li*, James C. Root, Thomas M. Atkinson, Tim A. Ahles

Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA

*Corresponding author at: Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, 641 Lexington Ave., 7th Floor, New York, NY 10022, USA. Tel.: +1-646-888-0047; fax: +1-212-888-2584. *E-mail address:* liy12@mskcc.org (Y. Li).

Abstract

Objective. Patient-reported cognition generally exhibits poor concordance with objectively assessed cognitive performance. In this article, we introduce latent regression Rasch modeling and provide a step-by-step tutorial for applying Rasch methods as an alternative to traditional correlation to better clarify the relationship of self-report and objective cognitive performance. An example analysis using these methods is also included.

Method. Introduction to latent regression Rasch modeling is provided together with a tutorial on implementing it using the JAGS programming language for the Bayesian posterior parameter estimates. In an example analysis, data from a longitudinal neuro-cognitive outcomes study of 132 breast cancer patients and 45 non-cancer matched controls that included self-report and objective performance measures pre- and post-treatment were analyzed using both conventional and latent regression Rasch model approaches.

Results. Consistent with previous research, conventional analysis and correlations between neurocognitive decline and self-reported problems were generally near zero. In contrast, application of latent regression Rasch modeling found statistically reliable associations between objective attention and processing speed measures with self-reported Attention and Memory scores.

Conclusions. Latent regression Rasch modeling, together with correlation of specific self-reported cognitive domains with neuro-cognitive measures, helps to clarify the relationship of self-report with objective performance. While the majority of patients attribute their cognitive difficulties to memory decline, the Rash modeling suggests the importance of processing speed and initial learning. To encourage the use of this method, a step-by-step guide and programming language for implementation is provided. Implications of this method in cognitive outcomes research are discussed.

Keywords: Statistical methods; Quality of life; Learning and memory

Introduction

Similar to findings in other cognitive syndromes (Marino et al., 2009; Middleton, Denney, Lynch, & Parmenter, 2006), previous research on the association of self-reported cognitive dysfunction in cancer survivors with objectively tested performance on neurocognitive measures suggests a weak to non-existent relationship. As an example, many cancer survivors treated with chemotherapy report difficulties in forgetfulness (Ahles et al., 2002; Cull, Stewart, & Altman, 1995; Cull et al., 1996; Schagen et al., 1999), which suggests significant disruption in memory at the core of mild cognitive impairments after cancer. However, despite the consistent report of memory dysfunction in survivors, research utilizing objective memory measures generally fails to find strong or consistent evidence of significantly affected memory performance post-treatment (Jenkins et al., 2006; Jim et al., 2012; Schagen, Muller, Boogerd, Mellenbergh, & van Dam, 2006; Wefel, Lenzi, Theriault, Davis, & Meyers, 2004).

While we cannot assume that individuals have perfectly accurate assessments of their own abilities, we believe one potential contributor to subjective–objective disagreement may be the limitations of conventional statistical methods that are used for

summarizing results of self-reported cognitive functioning. These methods rely on aggregating item responses into total scores or subdomains and submitting these to traditional, correlation-based analyses. While efficient, simply summing self-report item responses into global or domain scores obscures a wealth of information at the item response level that may lead to an inaccurate impression of subjective functioning and relies on untested and problematic assumptions that: (1) an increase in one point on one item is equivalent in its functional impact to an increase in one point on another item; (2) an endorsement of "somewhat often" on two individual items is considered equal, even if one item is relatively rare and the other relatively more commonly endorsed; and (3) global or domain level scores have the same sensitivity and specificity in their association with neurocognitive performance as item level scores.

The intent of this article is both as a "how-to" tutorial on using advanced statistical methods—latent regression Rasch modeling—to more accurately examine the association of subjective and objective performance to help researchers quickly and effectively apply these methods in their own research and move beyond the conventional correlation-based method, as well as to serve as an example of its application in assessing the association of subjective and objective reports of cognitive dysfunction in cancer survivors. Mathematics are kept at a minimum, and we describe only the essential equations, so that the analyses can be clearly explained. Statistical computer programs using a Bayesian Gibbs sampling approach are provided as recipes to facilitate these analyses, which are not overwhelmingly different from an ordinary linear regression model, a method already familiar to most researchers in our field. To complete the tutorial, example analyses are provided to illustrate the strengths and limitations of conventional and advanced methods as it is applied in clarifying the association of survivors' self-reported attention and memory dysfunction with objective neurocognitive assessment.

Latent Regression Rasch Model Principles and Tutorial

Latent regression Rasch modeling (De Boeck & Wilson, 2004) (henceforth latent Rasch for brevity) has advantages over a conventional psychometric approach (Bond & Fox, 2001; Embretson & Reise, 2000) described above. First, the latent Rasch approach directly models individual, item-level, cognitive symptom ratings, whereas the conventional approach aggregates over symptom ratings to form subscale or global scores, which inevitably obscures specific patterns of symptoms. As a result, while a conventional analysis makes no distinction between the sum of 1 and 4 or 2 and 3 on individual item responses, the latent Rasch approach does take advantage of this information. Secondly, the latent Rasch model weights endorsement of rare symptoms more highly than commonly reported symptoms, whereas the conventional approach normally weights all symptoms identically. Thirdly, the latent Rasch approach facilitates the development of a predictive tool simultaneously considering all available information, both the patient-reported subjective symptoms and objective neurocognitive measures.

Below is a brief introduction for practical work. As the name implies, it consists of two parts, a Rasch model and a regression model. We first cover the Rasch model, which is probably the less familiar part, and then we add the regression extension to complete the latent regression Rasch model. The Rasch model (Rasch, 1980) is a widely used model to score item responses, such as the raw 48-item MASQ data, into a single numeric number representing, in this case, each patient's self-reported cognitive problems and issues on a latent scale from no problem to numerous problems (a higher score represents more problems as per p. 95 of Seidenberg, Haltiner, Taylor, Hermann, & Wyler, 1994). For example, item 4 asks "I forget to mention important issues during conversations" and responses of "frequently" or "almost always" may be taken to represent strong endorsement of problems and issues (when compared with the remaining "almost never," "rarely," and "sometimes"). Twenty-three items require reverse scoring, for example, "After putting something away for safekeeping, I am able to recall its location"; responses of "sometimes," "rarely," or "almost never" represent the presence of problems.

The simplest Rasch model involves dichotomized item response data. Therefore, we dichotomize each item according to its direction of coding (higher scores always represent greater cognitive dysfunction), so that 1 = problem present (a response of "frequently" or "almost always") and 0 = no problem ("almost never," "rarely," and "sometimes"). A hypothetical rectangular data set is shown below:

$$
\begin{array}{c}
\quad\quad \beta_1 \;\; \beta_2 \;\; \beta_3 \;\; \ldots \;\; \beta_{48} \\
\begin{array}{c}
\theta_1 \\
\theta_2 \\
\theta_3 \\
\theta_4 \\
\vdots \\
\theta_{177}
\end{array}
\left[
\begin{array}{ccccc}
1 & 0 & 1 & \ldots & 1 \\
0 & 0 & 0 & \ldots & 1 \\
0 & 1 & 1 & \ldots & 1 \\
1 & 1 & 1 &   & 1 \\
  & \vdots & & \ddots & \vdots \\
0 & 1 & 1 & \ldots & 1
\end{array}
\right],
\end{array}
$$

where the columns represent the MASQ items and the rows represent respondents. Each item has a threshold parameter (β parameters, β_1, $\beta_{2,...,}\beta_{48}$). Each of the 177 hypothetical respondents has a latent score (θ parameters, θ_1, $\theta_{2,...,}\theta_{177}$), which represents each subject's

self-reported cognitive problems. The scale of the θ parameter is not directly observable, but its effects are manifested in the item responses and thus can be estimated.

A specific example helps illustrate the model. Suppose that a person has an underlying cognitive problem of 1.2 ($\theta_3 = 1.2$ in symbols). The absolute unit of the scale can be ignored for now. Also assume that items 1 and 2 have a severity threshold of 1.8 and 0.5, respectively ($\beta_1 = 1.8$, $\beta_2 = 0.5$). According to the Rasch model, the person is likely to report a problem in item 2 because her underlying performance of 1.2 *exceeds* the threshold for item 2. But she is unlikely to endorse item 1 because $\theta_3 = 1.2$ *falls below* the threshold of $\beta_1 = 1.8$ for item 1. We can work out the contrasts $\theta_3 - \beta_1$ and $\theta_3 - \beta_2$, which are -0.6 and $+0.7$, respectively. Therefore, a contrast between a person's underlying cognitive performance and an item's threshold, e.g., $\theta_3 - \beta_1$, represents the likelihood of a person endorsing item 1. Mathematically, the Rasch model is analogous to a logistic regression in the sense that the probability of endorsing an item, denoted as p_{ij}, is an inverse logit function of two predictors θ_i and β_j:

$$p_{ij} = \text{logit}^{-1}(\theta_i - \beta_j) = \frac{\exp(\theta_i - \beta_j)}{1 + \exp(\theta_i - \beta_j)}, \tag{1}$$

where the item responses are distilled into two sets of numbers—each person's underlying cognitive problem and each item's severity threshold. Equation (1) is the Rasch portion of the model.

Next, we turn to the latent regression part of the model. Each person's underlying cognitive performance, the θ_i estimates in Equation (1), can be further modeled by predictors in a regression. For example, we may hypothesize that cognitive performance at the 6 months post-treatment time point is associated with declines of a neurocognitive measure since baseline, which can be expressed as a predictor $\triangle x_i$, where $\triangle x_i$ may represent post-treatment declines in attention, memory, or any objectively measured neurocognitive performance. Each person's underlying self-reported cognitive problems θ_i is then regressed on $\triangle x_i$ as $\theta_i = \vartheta_0 + \vartheta_1 \cdot \triangle x_i + \varepsilon_i$, which follows the familiar form of a linear regression with an intercept ϑ_0 and a slope ϑ_1.

Merging this regression into the Rasch model, we have

$$\begin{cases} p_{ij} = \text{logit}^{-1}(\theta_i - \beta_j) \\ \theta_i = \vartheta_0 + \vartheta_1 \cdot \Delta x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2), \end{cases} \tag{2}$$

which completes the latent regression Rasch model. Of particular interest is the slope parameter, ϑ_1, whose value represents the extent to which post-treatment *declines* in the predictor affects post-treatment cognitive difficulties and problems. Equation (2) is essentially a set of simultaneous equations and can be further simplified into one line,

$$p_{ij} = \text{logit}^{-1}([\vartheta_0 + \vartheta_1 \cdot \Delta x_i + \varepsilon_i] - \beta_j), \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2), \tag{3}$$

by incorporating the second equation into the first, thereby simultaneously explaining item responses in MASQ at 6 months and declines in a neurocognitive test since baseline. Additional covariates can be included, such as cancer treatment to account for the differences in MASQ outcomes due to cancer treatment.

The main rationale for a lagged analysis is that it helps to achieve a clearer causal link and to minimize misleading interpretations of causal relationships in cross-sectional data analyses for behavioral research (Weinstein, 2007). If the outcomes of interest and its putative predictors are collected at the same assessment time point, then a cross-sectional correlation between the two can lead to an inflated estimate of their true association. A cross-sectional analysis cannot clearly tease out the level of inflation between neurocognitive performance and the subjective, self-reported MASQ score. A lagged analysis is less prone to this problem, such as regressing outcomes measured at a later time point by predictors at an earlier time point (or changes since).

Example Analysis Using Conventional and Latent Regression Rasch Model Approaches

In our example analysis using latent Rasch methods, we were specifically interested in: (i) comparing results from conventional versus Rasch methods on real data; and (ii) assessing the relationship of attention and memory complaints with objective learning and memory performance. For the second aim, previous research has typically assessed subjective–objective agreement using self-reported memory complaints and objective delayed memory performance. However, for the individual, the subjective experience of forgetfulness may not be due to true forgetting of information, but rather to inadequate attention and encoding of information at the time of learning. Previous work from our lab (Root, Andreotti, Tsu, Ellmore, & Ahles, 2015; Root, Ryan et al., 2015), which studied the learning and memory performance of breast cancer survivors on serial list learning tasks [California Verbal Learning Test-2 (CVLT-2); Brown Location Test], suggests that cancer survivors exhibit poor initial attention and encoding of information (Trial 1; Middle Region Recall) that is compensated for by repetition (Trial 5), yielding normal retention and

recall of this information at delay (Long Delay Free Recall). As a result, a focus on delayed memory performance with subjective report will underestimate subjective–objective agreement, while focusing on initial attention and encoding variables (Trial 1; Middle Region Recall) and their relation to subjective memory complaints may be a more relevant comparison. To the extent that subject-reported memory complaints are driven by suboptimal attentional performance, we were interested in applying the latent Rasch model to clarify the relationship of self-reported attention and memory function with objective performance on measures of attention, psychomotor speed, and memory. As a result, in addition to inclusion of Long Delay Recall performance, we also included attention and initial encoding variables previously identified (Trial 1; Middle Region Recall) together with measures of sustained attention and psychomotor speed, hypothesizing that these would be associated with self-reported attention and memory function.

Materials and Methods

Research Context

The latent Rasch model was applied to measures collected in a parent study (Ahles et al. 2010, 2014). Briefly, in a prospective longitudinal observational study on neurocognitive impairments after breast cancer treatment, we administered a battery of neurocognitive measures to patients exposed to chemotherapy ($n = 60$), patients not exposed to chemotherapy ($n = 72$), and non-cancer controls ($n = 45$). A battery of standardized neurocognitive measures were administered, and scored onto a z-scale so that one unit difference in the neurocognitive test scores reflects a 1 *SD* difference in the norm. Participants were also assessed by the Multiple Ability Self-Report Questionnaire (MASQ; Seidenberg et al., 1994). The MASQ contains 48 items covering 5 domains: Language, Visual-perceptual ability, Verbal Memory, Visual Memory, and Attention/Concentration. The assessments were made before treatment and at 1, 6, and 18 months after treatment (or at a yoked schedule for the non-cancer controls). The study was approved by appropriate Institutional Review Boards.

Analysis

The neurocognitive scores were longitudinal (pre-treatment to 6-month post-treatment or at yoked intervals for controls) and MASQ scores were cross-sectional at 6 months. For the present analysis, we aimed to relate MASQ Total score, Attention and Verbal Memory subscales at the 6-month time point to declines in tested neurocognitive performance from pre-treatment to post-treatment. The primary question we asked was whether post-treatment self-reported MASQ problems on attention and memory functioning were related to specific declines in objectively assessed neurocognitive performance.

Two sets of analyses were carried out. The first analysis was based on a conventional analytic approach, by calculating essentially a Pearson correlation matrix between the MASQ scores at 6 months post-treatment and pre- to post-treatment changes in neurocognitive performance assessed objectively. The second analysis was based on the latent Rasch approach, by running a series of 24 latent Rasch models as outlined above, each with the full MASQ responses of 48 items at the post-treatment time-point and change scores from pre-treatment to post-treatment in each neurocognitive performance variable. All processing steps follow the analysis steps described in the above tutorial.

The latent Rasch modeling by Markov Chains provide more information than what is available in conventional correlation analysis. The β_j parameters provide estimates of each symptom's threshold, with rarely endorsed items having a higher threshold in general. The θ_i parameters (each person's latent MASQ problems) are typically scaled to follow a standard normal distribution with a mean of 0 and a standard deviation of 1, thereby conveniently mapping all parameters onto the z-scale. The latent Rasch model scores each person's latent MASQ severity differently than conventional scoring. Conventional scoring makes no distinction between rare and common symptoms. For example, only 2.4% of respondents endorsed a problem in "It is easy for me to read and follow a newspaper story," whereas 26% of respondents identified a problem in "I find myself searching for the right word to express my thoughts," although in conventional scoring endorsement of either would yield an equivalent increase in MASQ score. In contrast, latent Rasch modeling would give a higher score to rarely endorsed symptoms than to commonly endorsed symptoms. Finally, Markov Chains yield the entire Bayesian posterior distribution for each parameter, which can be used to derive, for example, the probability that patients report more memory problems than attention problems on the MASQ.

Statistical Programming Syntax Using JAGS

We used the JAGS statistical software to estimate the parameters, with commonly used settings, e.g., running 3 Markov Chains with 2500 burn-ins discarded, a thinning interval of 10 (only sampling every 10th simulated iterations to reduce auto-correlation), and a final total of 5000 simulations saved which yielded estimates of the entire posterior distribution. The command syntax is

listed in the Appendix. Briefly, the essence of the commands is found in lines 5–10, where the model syntax closely resembles Equation (3). As the JAGS program calculates the Bayesian posterior distributions of the parameters, their prior distributions are required (lines 15–21, which show generally flat and thus uninformative priors). A more detailed summary of how to install JAGS and related packages can be found in the Supplementary material online. Tutorials on the JAGS language abound. A good starting point is in the examples that come with the downloaded JAGS software package.

Results

Conventional Correlational Analysis

Table 1 shows the Pearson correlation coefficients between the MASQ scores and several neurocognitive performance measures. As per scoring conventions, the MASQ global scale and the Attention and Verbal Memory subscales were coded, so that a higher score represented *greater* problems and issues. Thus, a negative correlation was the anticipated association, i.e., that a post-treatment *decline* in neurocognitive performance was associated with *greater* self-reported problems in the MASQ, and conversely that an *improvement* in post-treatment neurocognitive performance was associated with a *lower* MASQ score (meaning less problems and issues). Table 1 shows statistically reliable associations between MASQ scores and neurocognitive performance variables. Speeded Number Sequencing was associated with the MASQ total score as well as the Attention subscale. Letter Sequencing was associated with the MASQ total scale, Attention, and Verbal Memory subscales. Observed correlations were relatively low and between 0.20 and 0.30. No other association was found between the MASQ scales and the neurocognitive performance measures, with generally near-zero correlation coefficients in remaining measures. Table 2 summarizes the neurocognitive assessments and their abbreviations presented in this article.

Latent Regression Rasch Model Analysis

Fig. 1 provides a visual summary of a series of latent Rasch model. The topmost line segment in Fig. 1, for example, represents the extent to which the latent scale of patient-reported MASQ problems is associated with objectively assessed post-treatment neurocognitive declines in Sustained Attention-Vigilance. Equations (2) and (3) show that the MASQ scores and the neurocognitive assessments were all mapped onto the standard normal z-scale. Thus, in this specific example, one unit of post-treatment

Table 1. Pearson correlation coefficients between MASQ Total Score, Attention and Verbal Memory subscales and a selected list of neurocognitive measures ($n = 165$)

| | Mean | Stdev | MASQ | | |
			Total	AC	VM
MASQ Total Score and Subscales (Verbal Memory and Attention)					
MASQ Total	93.24	23.90			
MASQ Attention (AC)	15.64	4.61	0.892**		
MASQ Verbal Memory (VM)	17.16	5.28	0.901**	0.770**	
Post-treatment changes in neurocognitive tests					
CVLT-2-LDFR	0.28	0.92	−0.090	−0.089	−0.051
CVLT-2-Trial 1	−0.05	1.25	−0.006	−0.043	−0.034
CVLT-2-MRR	0.09	1.08	−0.082	−0.008	−0.016
CVLT-2-List B	−0.18	0.98	−0.052	−0.025	−0.053
PASAT 2″	0.34	0.71	−0.124	−0.087	−0.109
D-KEFS TMT-Visual Scanning	0.07	0.62	−0.049	0.006	−0.134
D-KEFS TMT-Number Sequencing	0.22	1.01	−0.200*	−0.195*	−0.131
D-KEFS TMT-Letter Sequencing	0.31	0.93	−0.278**	−0.244**	−0.286**
D-KEFS CWI-Color Naming	0.002	0.74	−0.003	−0.002	0.035
D-KEFS CWI-Word Reading	−0.13	0.60	−0.030	−0.042	−0.064
D-KEFS CWI-Inhibition	0.18	0.65	−0.055	−0.058	−0.080
CPT-Distractibility	0.05	0.87	0.006	0.020	0.041
CPT-Vigilance	0.16	0.65	−0.017	0.009	0.005

Notes: Listed are the Pearson correlation coefficients between the MASQ scores measured at 6 months post-treatment and the changes of neurocognitive performance between baseline and 6 months post treatment. The MASQ was scored so that a higher score represented worse outcomes (*greater* problems or issues). Thus, a negative correlation was the anticipated direction of correlation, meaning that a decline in neurocognitive performance at 6-months post-treatment was associated with more self-reported MASQ problems.
*$p < .05$; **$p < .01$.

Table 2. Summary of neuropsychological assessments and their abbreviations

Neuropsychological measure	Abbreviated
Reaction time	
Continuous Performance Test-Vigilance	CPT-Vigilance
Continuous Performance Test-Distractibility	CPT-Distractibility
Processing speed	
D-KEFS Color Word Interference Test-Inhibition/Switching	D-KEFS CWI-Inhibition
D-KEFS Color Word Interference Test-Word Reading	D-KEFS CWI-Word Reading
D-KEFS Color Word Interference Test-Color Naming	D-KEFS CWI-Color Naming
D-KEFS Trail Making Test-Motor Speed	D-KEFS TMT-Motor Speed
D-KEFS Trail Making Test-Number-Letter Switching	D-KEFS TMT-N/L Switching
D-KEFS Trail Making Test-Letter Sequencing	D-KEFS TMT-Letter Sequencing
D-KEFS Trail Making Test-Number Sequencing	D-KEFS TMT-Number Sequencing
D-KEFS Trail Making Test-Visual Scanning	D-KEFS TMT-Visual Scanning
Grooved Pegboard (Right)	Pegboard (R)
Grooved Pegboard (Left)	Pegboard (L)
WAIS-III Digit Symbol Coding	WAIS-III Digit Symbol
Working Memory	
Rao Paced Auditory Serial Addition Test 3″	PASAT 3″
Rao Paced Auditory Serial Addition Test 2″	PASAT 2″
Visual Memory	
WMS-III-Faces I	WMS-III-Faces I
WMS-III-Faces II	WMS-III-Faces II
Verbal Memory	
CVLT-2-Trial 1	CVLT-2-Trial 1
CVLT-2-Middle Region Recall	CVLT-2-MRR
CVLT-2-List B	CVLT-2-List B
CVLT-2 Trials 1–5 Total	CVLT-2-Trials 1–5
CVLT-2-Long Delay Free Recall	CVLT-2-LDFR
Verbal Fluency	
D-KEFS-Letter Fluency	D-KEFS VFT-Letter
D-KEFS-Category Fluency	D-KEFS VFT-Category

decline in Sustained Attention-Vigilance was associated with *greater* MASQ problems by 0.295 in *z*-scale, with a 90% posterior Highest Density Region (0.011, 0.585), excluding the null (HDRs, see Gelman, Carlin, Stern, & Rubin, 2003). The 95% posterior HDR (-0.045, 0.640) failed to exclude the null. Fig. 1 summarizes the regression slope estimates (ϑ_1 parameter in the latent Rasch model), i.e., the extent to which post-treatment declines in neurocognitive performance affect the latent MASQ problems measured post-treatment. Neurocognitive performance measures that made no statistically reliable contribution were plotted in gray to simplify the plot. Six neurocognitive tests were reliably associated with the MASQ total score, including the three in Table 1. Three additional tests, shown in Table 1 with near-zero correlations calculated by conventional analysis, were found to be reliably associated with the outcome by latent Rasch.

Fig. 1 shows variability in how objectively assessed neurocognitive performance affected the latent MASQ Total symptoms. For example, within our Verbal Memory domain, only the Percent Middle-Region Recall, considered a measure of active attention based on previous research, had a statistically reliable effect of 0.240 (90% posterior HDR: 0.057, 0.424), indicating that decline in Middle-Region Recall was associated with greater self-reported problems on the MASQ. Remaining verbal memory variables (Trial 1, List B, CVLT Total, and Long Delay) had no significant associations. Outside of verbal memory variables, using the 90% posterior HDRs as a guide, neurocognitive measures reliably associated with global MASQ problems included Sustained Attention-Vigilance, Speeded Line Tracing, Speeded Letter and Number Sequencing, and Speeded Visual Scanning. All these tests had the same pattern of association—post-treatment declines were reliably associated with greater self-reported difficulties on the MASQ. Global MASQ symptoms were not associated with tests in the domains of Verbal Fluency, Visual Memory, or Working Memory since none of the tests in these domains had a 90% posterior HDR that excluded the null.

Association of Post-treatment, Domain-Specific MASQ Scores with Neurocognitive Variables

Fig. 2 plots the extent to which each of the neurocognitive variables affected the MASQ Attention and Verbal Memory subscales. For the MASQ Verbal Memory Subscale, subjective verbal memory abilities were associated with changes in psychomotor
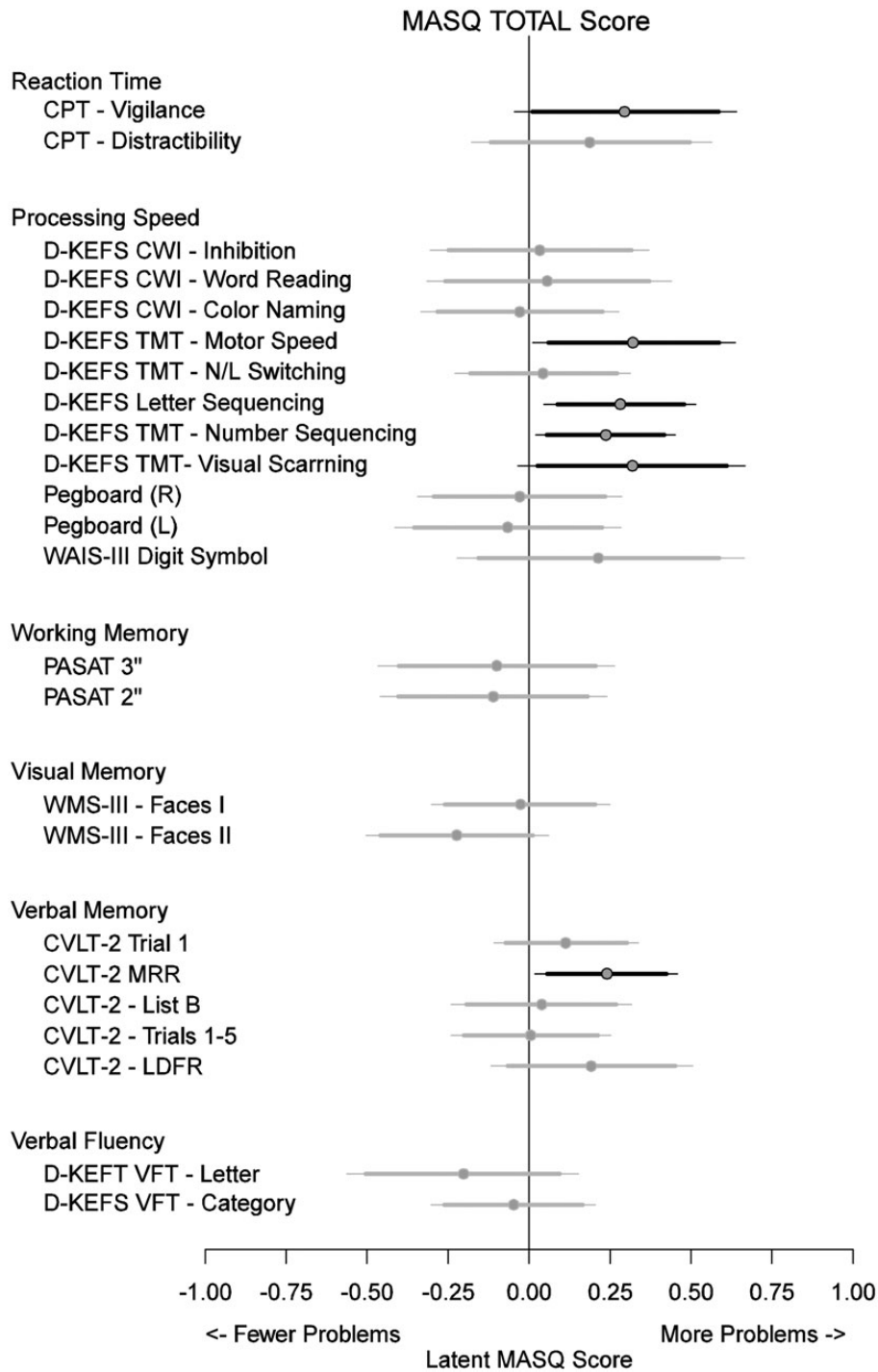
Fig. 1. Plot of associations between longitudinal changes in neurocognitive tests and patient-reported problems and difficulties in post-treatment MASQ total score. Circles represent the posterior means of the extent to which, for example, declines in post-treatment Sustained Attention is associated with greater MASQ symptoms at 6 months. The line segments represent the posterior highest density regions, with the thick and thin lines representing the 90% and 95% coverage, respectively. Plotted in gray are neurocognitive tests not reliably different from zero at the 90% level.

speed (Speeded Line Tracing; Speeded Letter Sequencing) as well as changes in attention and recall (Percent Middle Region Recall; Sustained Attention-Vigilance; Long Delay Free Recall). Subjective assessment of attention was associated with changes in psychomotor speed (Speeded Letter Sequencing; Speeded Number Sequencing; Speeded Visual Scanning), semantic fluency (Category Fluency), and attention (Trial 1; Percent Middle Region Recall, Sustained Attention-Vigilance).
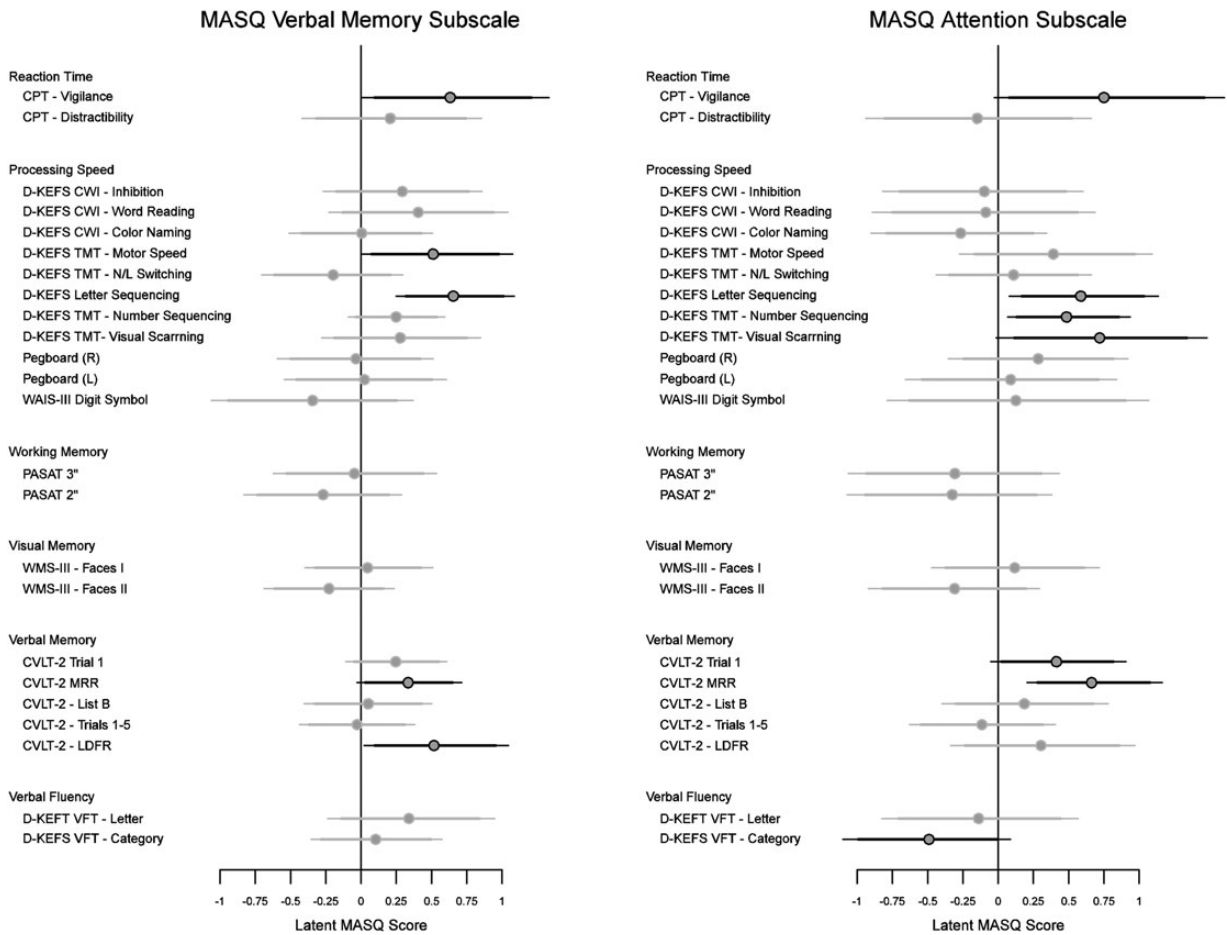
Fig. 2. Plots of highest density regions on the effects of neurocognitive tests on the Verbal Memory and Attention subscales of MASQ. Lines are plotted in gray if their corresponding 90% highest density regions fail to exclude the null. More problems are expected in Attention if there is a post-treatment reduction in the Trail-Making trials. Overall, Trail-Making measures and Percent Recall in Middle Region are consistently associated with both subscales.

Visual Exploration of Latent Regression Rasch

Fig. 3 shows how latent Rasch operates in comparison with a conventional correlation. Fig. 3a plots the Rasch-estimated, total latent cognitive problems as a function of observed changes in CVLT Middle-Region Recall. Each circle represents one person. The circles are color-coded according to the total number of MASQ symptoms.

Fig. 3 requires three pieces of information. The first is the change in CVLT Middle-Region Recall score from baseline to post-treatment. The second is the total number of patient-reported MASQ symptoms at post-treatment. The third is the Rasch-estimated latent cognitive problem. The first two variables are already available in the data set. The Rasch-estimated latent score can be obtained from the JAGS output, as illustrated in the Supplementary material online. The JAGS output in that example includes the mean parameter estimates labeled "theta[1]", "theta[2]", etc., which represents the Rasch-estimated latent cognitive problems for persons 1, 2, and so on. They can be collected from the JAGS output and plotted against the observed changes in the CVLT-II Middle Region Recall scores, which produces Fig. 3a. Then, we color code the circles in the plot to represent the total number of MASQ symptoms. For instance, the circle representing participant 51 is marked in Fig. 3a by an arrow, near an *x*-value of 2.0 and a *y*-value of 0.50. Her CVLT Middle-Region Recall score improved by exactly two points, and she reported a total of 4 MASQ symptoms and thus was plotted in the color indicated in the figure legend. Her Rasch-estimated latent score, obtained from the JAGS output, was 0.56 (more precisely, 0.560094). This process is repeated for all participants to complete Fig. 3a. Slight variations are visible in the latent Rasch scoring in Fig. 3a. In contrast, no variations are available in the conventional scoring of summing over endorsed items in Fig. 3b. By way of illustration, the latent Rasch scoring is accompanied by a slightly higher regression slope than the conventional scoring.
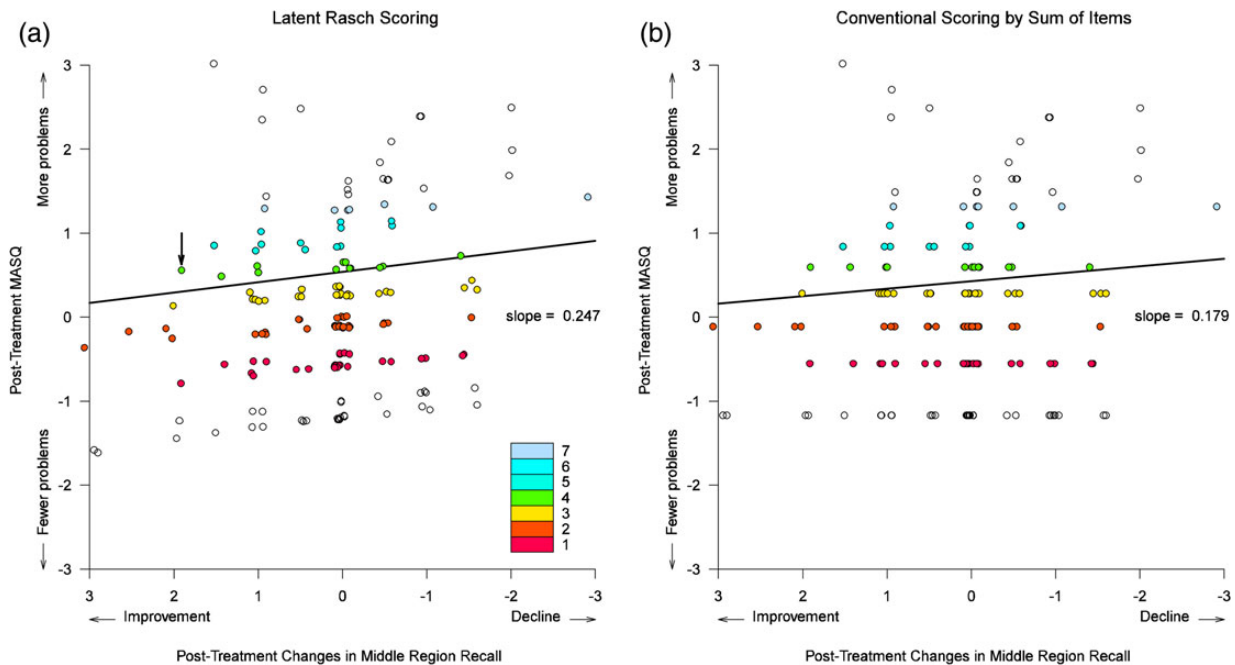
Fig. 3. Visual comparison between Rasch scoring and conventional scoring of total MASQ neurocognitive problems. Plotted in (a) are the Rasch-estimated latent neurocognitive issues against the observed changes in CVLT Middle-Region Recall. Plotted in (b) are what conventional correlations would produce by summing item scores. The small circles are color-coded to represent the same total number of symptoms. Random perturbations were added to the horizontal positions alone to visually separate the circles. Rasch scoring shows small variations in the total latent neurocognitive problems, whereas the conventional scoring assigns identical scores by summing the total endorsed problems. Rasch scoring was associated with a slightly higher regression slope than would conventional scoring.

Predictions Made by the Latent Regression Rasch Model

Fig. 4 plots the model-predicted MASQ subscale responses. We took the model-estimated number of symptoms in the five MASQ subscales and plotted them as five curves: Language, Visual Perceptual, Verbal Memory, Visual Memory, and Attention, over a range of hypothetical latent neurocognitive problems from -3 to $+3$. The norm (latent $\theta_i = 0.0$ on the x-axis) would be associated with an expected symptom of nearly 7 out of the total 8 in the MASQ Verbal Memory subscale. The expected number of Verbal Memory symptoms would rise quickly to 8 for someone whose latent neurocognitive problems are 2 *SD* above the norm. The expected number of Verbal Memory symptoms would drop to 3 for someone whose latent neurocognitive problems are at 2 *SD* below the norm.

The curve representing Verbal Memory is highest in the plot. The latent Rasch model predicted that patients would attribute their cognitive problems as primarily problems in Verbal Memory (e.g., "I have to hear or read something several times before I can recall it without difficulty"). In contrast, patients were least likely to report impairments in Visual-Spatial Memory (e.g., "After putting something away for safekeeping, I am able to recall its location") and in Attention (e.g., "I am easily distracted from my work by things going on around me"). The latent Rasch predicts that Attention would be the second least-reported subscale.

We estimated, using the Bayesian Markov-Chain simulations saved in JAGS, the probability that patients would report greater symptoms in Verbal Memory than in Attention. We found that, over 99.9% of the time, patients would report more Verbal Memory problems than Attention problems.

Discussion

Conventional analytics based on correlation coefficients show mostly null associations between measures of objective neurocognitive performance and self-reported cognitive problems, with nearly all correlations no greater than 0.20—a weak and mostly non-significant association. Exceptions are found, in the 0.20–0.30 range of association, in only two Trail Making measures, where statistically reliable correlations are observed in Speeded Letter Sequencing and Speeded Number Sequencing. A conventional statistical approach based on correlation coefficients would conclude that the self-reported, post-treatment cognitive problems are weakly associated with changes in Processing Speed, specifically in letter and number sequencing tests. All other
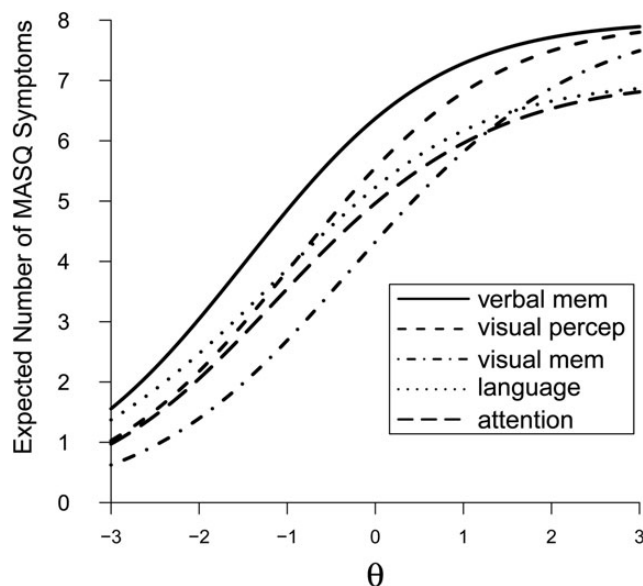
Fig. 4. Model-estimated number of symptoms plotted as a function of hypothetical latent cognitive problems. The five curves represent the five MASQ subscales: Language, Visual Perceptual, Verbal Memory, Visual Memory, and Attention. There are 8 items in four of the five subscales (the exception, Visual Perceptual subscale, has 6). The maximum value on the *y*-axis is plotted at 8 to facilitate interpretation. The expected Visual Perceptual symptoms are rescaled to have a maximum of 8. The curve representing Verbal Memory has the greatest expected number of symptoms than all the other subscales, indicating that problems in Verbal Memory are the most prevalent problems reported by the patients across the full range of hypothetical latent problems. In contrast, Visual Memory issues are the least reported problems. Attention is the second least-reported subscale of problems. The model shows that patients attribute their cognitive problems as primarily issues in Verbal Memory and not as issues in Attention and Visual Memory.

neurocognitive domains, Working Memory, Visual Memory, Verbal Fluency, Verbal Memory, Distractibility, and Reaction Time, would not appear to be associated with self-reported cognitive functioning using conventional approaches alone.

The use of the latent Rasch model clarifies and extends the pattern of associations between self-report and objective performance. By going beyond the conventional correlation-based approach, we have observed consistent patterns between the two sets of measurements. In addition to replicating the significant contributions in Speeded Letter Sequencing and Speeded Number Sequencing tests, we also found that: (i) subjective memory performance was associated with changes in delayed recall performance (Long Delay Recall), with changes in attention and encoding (Sustained Attention, CVLT Middle Region Recall), and in processing speed (Speeded Line Tracing and Number Sequencing) (Fig. 2); and (ii) subjective attention performance was associated with changes in attention and encoding (Sustained Attention, CVLT Trial 1, and CVLT Middle Region Recall), and with changes in processing speed (Speeded Letter Sequencing, Number Sequencing, and Visual Scanning) (Fig. 2). This distinction is found using the latent Rasch approach; however, the same pattern fails to appear in conventional analytics. These observations have methodological, clinical, and conceptual implications.

Methodologically, with regard to the superiority of the latent Rasch model to more conventional approaches, we note that these associations were not found in the more traditional, aggregate-domain level analysis. Methodologically, the discrepancy between the two statistical approaches is possibly due to the inadequacies of conventional analytic tools. Conventional correlation analysis weighs each MASQ item equally, while latent Rasch differentiates them. The difference between conventional and Rasch modeling can be seen in Fig. 3, which reveals the increased variation of self-reported cognitive problems using the latent Rasch model as opposed to more conventional summing of item responses. During conventional scoring, the raw responses are aggregated, so that variations between symptoms are lost in the final correlation analysis. The latent Rasch model extracts information directly from the raw item responses, allowing variations between symptoms to be part of the model, which results in a more streamlined and sensitive model.

Clinically, the model-estimated effect sizes are generally subtle, at a range of 0.25–0.30 in *z*-scores. But what does a subtle effect size of this magnitude mean in a clinical context? We can approach this by leveraging the Rasch-estimates. Fig. 1, for example, shows that a 1 *SD* decline in CVLT Middle-Region Recall is associated with a greater number of self-reported cognitive problems by a 0.24 *SD* above the MASQ norm. By further consulting Fig. 4, this above-norm 0.24 effect size would translate to, approximately, the endorsement of nearly 7 symptoms out of the maximum of 8 in the MASQ Verbal Memory subscale. Moreover, if we consider the margin of error of the effect size, at a posterior 95% highest density region of (0.02, 0.46), then we note that the

number of Verbal Memory complaints could be between 6.5 and 7.5 out of the maximum of 8. The latent regression Rasch model provides highly specific information beyond what can be culled from a conventional correlation analysis.

With regard to agreement between subjective and objective assessments, our results suggest that in fact subjective and objective assessments do generally agree in that subjective attention and memory difficulties are related to objective attention and memory performance, respectively. Results of our analysis additionally suggest that subjective memory complaints are also related to objective changes in attention and psychomotor speed, and subjective attention complaints are additionally related to objective changes not only in attention but also in processing speed.

Predictions made by the latent Rasch model also provide information not available in conventional correlation analysis. The predicted symptoms shed light on the discrepancy between subjective and objective measures of neurocognitive performance. The norm of underlying neurocognitive performance would be expected to bring about ∼6–7 symptoms of the maximum 8 Verbal Memory problems in MASQ, while in Attention only 5 symptoms. Fig. 4 shows a consistent pattern of patients attributing their problems more to forgetfulness than to deficits in learning, attention, and executive functioning. This gap reduces slightly as the underlying neurocognitive problems drop below the norm, suggestive of patients' misattribution of memory problems as their underlying cognitive difficulty worsens. The latent Rasch model allows the researcher to examine where the differences may be found. Aggregation of individual item responses to generate global or domain scores, as is typically done in a conventional analysis, may not identify these nuanced and subtle differences. The possibility of patients' misattribution may have implications in cognitive rehabilitation, where the conventional correlation analysis would suggest training in executive functioning and motor speed. However, the latent Rasch modeling would suggest training in attention, vigilance, and learning.

With regard to limitations, the primary goal of this article is to introduce readers to a novel application of latent regression Rasch modeling in assessing agreement between subjective and objective measures of neurocognitive impairment. The example analysis, described above, is included to demonstrate the performance of latent Rasch in comparison with more conventional approaches in clarifying the relationship of subjective report with objective performance. Group status is controlled for in our model, but is not reported in the findings as a contrast between groups. As a tutorial, the current example aims to motivate an understanding of what latent Rasch can do, that it can be implemented with relative ease, and it yields useful information not available in conventional correlations. However, users should be aware of model assumptions and potential pitfalls. Computationally, like many other latent models (Li & Baser, 2012), the latent regression Rasch model is not identified, that is, a constant can be added to the model parameters without affecting the model's prediction. An adjustment may be needed (i.e., the rescaling in the small hypothetical example in the Supplementary material online). The model also requires a local independence assumption, which states, in this specific context, that the endorsement of individual MASQ symptoms is mutually independent after the symptom endorsement threshold parameters are accounted for in the Rasch model. However, it is worth noting that this assumption is shared by all Rasch-like models (Embretson & Reise, 2000), a price to pay for accessing the item responses directly. The statistical assumptions notwithstanding, a model is likely to be welcomed and widely used if it is not difficult to apply and it yields useful information in addressing real research questions. The ultimate test of the latent regression Rasch model in explaining the agreement between objective and subjective neurocognitive assessments has to come from researchers applying it in addressing substantive questions.

The numerous neurocognitive models fitted may be a cause for concern for some researchers who may therefore question the validity of the present findings, i.e., whether some of the relationships between neurocognitive measures and MASQ symptoms represent chance findings, given the number of models. However, our latent Rasch modeling approach is no more susceptible to this concern than conventional analyses involving large correlation matrices, and we note that only few significant associations were found using the more traditional, correlation matrix approach. We note that one association, between subjective attention and semantic fluency performance, was the opposite of what would be expected (i.e., objective decline in performance on the Category Fluency measure was associated with lesser report of attention dysfunction). There is no clear, intuitive explanation for this association and this finding may represent an artifact of the number of comparisons included across all three Global, Memory, and Attention models. One method that would limit potentially false-positive findings when testing for associations between multiple measures consists of aggregation of neurocognitive measures (e.g., by combining all Trail Making tests into one). Another would be to limit the number of measures included to only those for which there are firm *a priori* predictions of association between objective and subjective measures. Since this was the first use of this statistical methodology to be applied to the problem of subjective–objective agreement, we chose to include all measures available to assess not only sensitivity to detect associations but also specificity. In this regard, the methodology works well, since with the exception of the verbal fluency finding, remaining associations found using the Rasch model are generally logical and supportable from a neurocognitive standpoint (i.e., changes in attention and psychomotor speed drive self-reported attention and memory complaints). As subjects' abilities to attend and more quickly and efficiently process information decline, so too does their subjective sense of their own attention and memory function.

In terms of the issue of false discovery, our Bayesian posterior HDRs claim as much validity as the data provide. A researcher can expand the posterior density regions from 95% to 99% or greater if he/she wishes to increase the confidence in the findings. Bayesians may disagree on the desired precision; however, they should all agree on how the posterior distributions are derived. As data accrue, these posterior intervals are likely to evolve to provide better precision and calibration. In the future, a more complex model may be built (e.g., Belin & Rubin, 1995), such as a hierarchical model that includes neurocognitive predictors nested within domains of deficits. Such a model, combined with general advantages of a Bayesian approach, would provide more clarity in the future.

An important aim of this paper is to introduce latent Rasch model to other researchers who previously relied on correlation analysis and its derivatives. The model is no more complex than ordinary regression, the JAGS syntax makes it straightforward to deploy, and it yields results and predictions not available in a conventional analysis. We hope that latent Rasch continues to be a useful tool to increasingly more researchers to further illustrate the possible mechanisms of cognitive impairment in cancer survivors.

Supplementary Material

Supplementary material is available at *Archives of Clinical Neuropsychology* online.

Funding

Conflict of Interest

None declared.

Appendix

```
1       model
        {
        for (p in 1:N) {
           for (i in 1:I) {
5       Y[p, i] ~ dbern( pr[p, i] )
           logit( pr[p, i] ) <- theta[p] - beta[i]
           }
           theta [p] <- Ychemo*chemo [p] + Nchemo*Nochemo [p] + bx*x [p] + eps [p]
           eps[p] ~ dnorm(0.0, tau)
10      }
        for (i in 1:I) {
           beta [i]  ~ dnorm(0.0, 0.001); b [i] <- beta [i] - mean (beta[])
           }
15      # Priors
        sigma ~ dunif(0, 10)
        tau <- 1 / (sigma * sigma)
        Ychemo ~ dnorm(0.0, 0.01)
20      Nchemo ~ dnorm(0.0, 0.01)
        bx ~ dnorm(0.0, 0.01)
        }
```

References

Ahles, T. A., Li, Y., McDonald, B. C., Schwartz, G. N., Kaufman, P. A., Tsongalis, G. J., et al. (2014). Longitudinal assessment of cognitive changes associated with adjuvant treatment for breast cancer: The impact of APOE and smoking. *Psychooncology*, *23 (12)*, 1382–1390.

Ahles, T. A., Saykin, A. J., Furstenberg, C. T., Cole, B., Mott, L. A., Skalla, K., et al. (2002). Neuropsychologic impact of standard-dose systemic chemotherapy in long-term survivors of breast cancer and lymphoma. *Journal of Clinical Oncology*, *20 (2)*, 485–493.

Ahles, T. A., Saykin, A. J., McDonald, B. C., Li, Y., Furstenberg, C. T., Hanscom, B. S., et al. (2010). Longitudinal assessment of cognitive changes associated with adjuvant treatment for breast cancer: Impact of age and cognitive reserve. *Journal of Clinical Oncology*, *28 (29)*, 4434–4440.

Belin, T. R., & Rubin, D. B. (1995). The analysis of repeated-measures data on schizophrenic reaction-times using mixture-models. *Statistics in Medicine*, *14 (8)*, 747–768.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cull, A., Hay, C., Love, S. B., Mackie, M., Smets, E., & Stewart, M. (1996). What do cancer patients mean when they complain of concentration and memory problems? *British Journal of Cancer*, *74 (10)*, 1674–1679.

Cull, A., Stewart, M., & Altman, D. G. (1995). Assessment of and intervention for psychosocial problems in routine oncology practice. *British Journal of Cancer*, *72 (1)*, 229–235.

De Boeck, P., & Wilson, M. (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York: Spinger-Verlag.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: LEA.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). *Bayesian data analysis*. New York: Chapman & Hall.

Jenkins, V., Shilling, V., Deutsch, G., Bloomfield, D., Morris, R., Allan, S., et al. (2006). A 3-year prospective study of the effects of adjuvant treatments on cognition in women with early stage breast cancer. *British Journal of Cancer*, *94 (6)*, 828–834.

Jim, H. S., Phillips, K. M., Chait, S., Faul, L. A., Popa, M. A., Lee, Y. H., et al. (2012). Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. *Journal of Clinical Oncology*, *30 (29)*, 3578–3587.

Li, Y., & Baser, R. (2012). Using R and WinBUGS to fit a generalized partial credit model for developing and evaluating patient-reported outcomes assessments. *Statistics in Medicine*, *31 (18)*, 2010–2026.

Marino, S. E., Meador, K. J., Loring, D. W., Okun, M. S., Fernandez, H. H., Fessler, A. J., et al. (2009). Subjective perception of cognition is related to mood and not performance. *Epilepsy and Behavior*, *14 (3)*, 459–464.

Middleton, L. S., Denney, D. R., Lynch, S. G., & Parmenter, B. (2006). The relationship between perceived and objective cognitive functioning in multiple sclerosis. *Archives in Clinical Neuropsychology*, *21 (5)*, 487–494.

Rasch, G. (1980). *Probabilistic models for some intelligency and attainment tests*. Chicago: The University of Chicago Press.

Root, J. C., Andreotti, C., Tsu, L., Ellmore, T. M., & Ahles, T. A. (2015). Learning and memory performance in breast cancer survivors 2 to 6 years post-treatment: The role of encoding versus forgetting. *Journal of Cancer Survivorship*, doi:10.1007/s11764-015-0505-4

Root, J. C., Ryan, E., Barnett, G., Andreotti, C., Bolutayo, K., & Ahles, T. (2015). Learning and memory performance in a cohort of clinically referred breast cancer survivors: The role of attention versus forgetting in patient-reported memory complaints. *Psycho-Oncology*, *24 (5)*, 548–555.

Schagen, S. B., Muller, M. J., Boogerd, W., Mellenbergh, G. J., & van Dam, F. S. (2006). Change in cognitive function after chemotherapy: A prospective longitudinal study in breast cancer patients. *Journal of National Cancer Institute*, *98 (23)*, 1742–1745.

Schagen, S. B., van Dam, F. S., Muller, M. J., Boogerd, W., Lindeboom, J., & Bruning, P. F. (1999). Cognitive deficits after postoperative adjuvant chemotherapy for breast carcinoma. *Cancer*, *85 (3)*, 640–650.

Seidenberg, M., Haltiner, A., Taylor, M. A., Hermann, B. B., & Wyler, A. (1994). Development and validation of a Multiple Ability Self-Report Questionnaire. *Journal of Clinical and Experimental Neuropsychology*, *16 (1)*, 93–104.

Wefel, J. S., Lenzi, R., Theriault, R. L., Davis, R. N., & Meyers, C. A. (2004). The cognitive sequelae of standard-dose adjuvant chemotherapy in women with breast carcinoma: Results of a prospective, randomized, longitudinal trial. *Cancer*, *100 (11)*, 2292–2299.

Weinstein, N. D. (2007). Misleading tests of health behavior theories. *Annals of Behavioral Medicine*, *33 (1)*, 1–10.