Title
Strategies for Analyzing Ordinal Quality-of-Life Data with Application to Patient's Assessment of Own Functioning Inventory

Author
Fedortsova, Nadejda Sergeevna

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Strategies for Analyzing Ordinal Quality-of-Life Data with Application to Patient's Assessment of Own Functioning Inventory

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Public Health

by

Nadejda Sergeevna Fedortsova

2020

ABSTRACT OF THE DISSERTATION

Strategies for Analyzing Ordinal Quality-of-Life Data
with Application to Patient's Assessment of Own Functioning Inventory

by

Nadejda Sergeevna Fedortsova
Doctor of Public Health
University of California, Los Angeles, 2020
Professor Thomas R. Belin, Chair

In quality-of-life research, the Patient's Assessment of Own Functioning Inventory (PAOFI) was developed by neuropsychologists to reflect capabilities in memory, language and communication, higher-level cognition, and sensorimotor functioning. A statistical perspective on applied research using the PAOFI raises questions about how ordinal item scores have been dichotomized into binary outcomes given what is known about the loss of information from dichotomizing continuously-scaled measures.

Drawing on a sample of breast-cancer survivors in a study of how breast-cancer treatment affected quality of life, and using an information-theory-based aggregate measure of entropy obtained by summing contributions of individual items, the score based on dichotomized PAOFI items was about 70% lower than using ordinal PAOFI items. Furthermore, investigation of PAOFI domain scores across breast cancer treatment groups revealed sensitivity of inferences to dichotomization cut-points, suggesting that avoiding dichotomization and analyzing PAOFI scores on the original ordinal scale might be preferable.

Previous investigation of the PAOFI on a diverse sample of individuals identified 4 promi-

nent factors; however, in the breast-cancer-survivor sample, a 5-factor solution provided a more natural interpretation. Analyses of domain scores across breast cancer treatment groups revealed significant differences in factor scores with some degree of sensitivity to whether the factor analysis used ordinal or dichotomized items. After using item response theory to select 2 items per domain, it was still possible to detect significant differences in domain scores across breast cancer treatment groups.

To reflect the full range of associations between background characteristics and quality-of-life domains, we implemented 3-part models for PAOFI outcomes, modeling (1) an indicator for any serious problem versus no serious problem within a given domain, (2) the number of problems experienced by a patient on a given domain, and (3) either the average severity or the aggregate impact of problems experienced on a given domain. Using the "Memory: Absent-Mindedness" domain for illustration, we found diverse sets of significant predictors of the different outcomes. Overall, the dissertation reveals a number of ways to improve upon previous statistical analysis approaches that dichotomized ordinal PAOFI items to enhance understanding of breast-cancer quality of life.

The dissertation of Nadejda Sergeevna Fedortsova is approved.

Abdelmonem A. Afifi

Patricia A. Ganz

David W. Gjertson

Donatello Telesca

Thomas R. Belin, Committee Chair

University of California, Los Angeles

2020

This work is dedicated to my parents

with deepest love and gratitude

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to my advisor Professor Thomas R. Belin for his continuous support, unending patience, and generosity with which he shared his knowledge and time with me. His constant encouragement, sense of humor, and empathy were vital in making this dissertation possible. I am beyond fortunate to have him as my "academic father".

I am deeply thankful to my committee members: Professor Abdelmonem A. Afifi, Professor Patricia A. Ganz, Professor David W. Gjertson, and Professor Donatello Telesca for their insightful comments, encouragement and support. A special thank you to Professor Patricia A. Ganz for letting me use the data from her study.

My appreciation also goes to UCLA Department of Biostatistics for their help in seeing this work come to fruition.

I am forever grateful to my mom, Dirk, Kitty, and Mason for their boundless love and support. Last but not least, I would like to thank my dad for instilling interest in math and science in me, and whom I wish stayed in this world long enough to know that I persevered.

VITA

2004	B.S. (Applied Mathematics), CSUN, Northridge, CA.
2009	M.S. (Biostatistics), UCLA, Los Angeles, CA.
2010–2016	Teaching Assistant, Department of Biostatistics, UCLA.
2012–2015	Project Leader, Baxter BioScience, Los Angeles, CA.
2015–2016	Research Consultant, National Health Foundation, Los Angeles, CA.
2016–2020	Graduate Student Researcher, Department of Biostatistics, UCLA.
2018	Lecturer, Division of Applied and Advanced Studies in Education, CSULA, Los Angeles, CA.
2018–2020	Lecturer, Department of Health Sciences, CSUN.

CHAPTER 1

Introduction

In research on quality-of-life for cancer survivors, cognitive functioning is an important consideration, such as in research on the so-called"chemo-brain" phenomenon where chemotherapy patients don't feel as mentally sharp after treatment as before. However, measuring one's cognitive wellness can be difficult. Part of the issue is that assessing cognitive function typically depends on patient self-reports, which often is done through multi-item questionnaires that, depending on their length, might be broken down into multiple subscales that can be viewed as distinct variables.

An example, viewed in work by Ganz et al. (2013), is the Patient's Assessment of Own Functioning Inventory (PAOFI), which was developed by neuropsychologists to reflect on patient capabilities in memory, language and communication, higher-level cognition, and sensorimotor functioning. The PAOFI consists of 33 items, each giving rise to an ordinal measurement on a 6-point scale, with previous work suggesting grouping these items into 4 subscales (memory, language and communication, higher-level cognition, and sensorimotor) (Chelune et al., 1986).

However, reviewing the use of the PAOFI from a statistical-analysis perspective gives rise to questions about how it has typically been used in practice. An approach that has been used in neuropsychology (e.g., Chelune et al., 1986; Richardson-Vejlgaard et al., 2009) involves dichotomizing the 6-point scale into binary outcomes, for example whether the assessment score was in $\{1, 2, 3\}$ or in $\{4, 5, 6\}$. As will be reviewed at greater length in this dissertation proposal, it is well known in statistics that dichotomizing measures on an

underlying continuous scale results in a loss of information, giving rise to concern that the same problem would adversely affect analyses of 6-point ordinal outcome data. A central goal of this dissertation work is to review and build upon the existing strategies for analyzing ordinally-scaled quality-of-life outcomes, both from the vantage point of achieving desirable precision and power of statistical tests, and from the vantage point of facilitating interpretation.

The broad outline of the dissertation proceeds in following fashion. First, we introduce the dataset from the motivating application and review the literature on loss of information due to collapsing more detailed measures into a smaller number of categories; we also review the literature on general analysis strategies for ordinal outcomes. We then explore a range of analysis methods on data provided by Dr. Patricia Ganz from Mind-Body Study, a prospective observational study of breast-cancer patients who have undergone a primary treatment, but not endocrine therapy, focusing on comparison of treatment exposures (chemotherapy only, radiation only, both chemotherapy and radiation, neither chemotherapy nor radiation) on various quality-of-life outcomes including the PAOFI. Because the PAOFI involves summing dichotomized variables, which can be expected to be affected by Central Limit Theorem patterns, we propose simulation experiments to investigate the extent to which outcome subscales comprised of dichotomized ordinal items affect significance findings. Subsequently, we propose a line of research to be undertaken as a dissertation project, building on existing findings and making use of ideas, such as multi-part models, which are frequently used in quality-of-life analyses and other health-services applications.

The rest of the dissertation is organized as follows. Chapter 2 contains the Mind-Body Study description, literature review of some methods often used for analysis of ordinally-scaled questionnaire data, and insights from the literature on information theory relevant to impact of dichotomizing measures on an underlying continuous scale. Chapter 3 reviews an investigation of different cut-points for PAOFI items and contrasts dichotomization with analyses carried out on an ordinal scale. Chapter 4 presents investigation of the factor-

analytic structure of the PAOFI data, with illustration of the effects of dichotomizing PAOFI items prior to conducting factor analysis on the composition of the resultant factors, as well as an exploration of the idea of a reduced version of the instrument using methodology of item response theory. Chapter 5 contains results from a three-part model fitted to the PAOFI data from the Mind-Body Study, with a focus on the role of different covariates in predicting distinct aspects of the PAOFI outcomes.

CHAPTER 2

Background and Literature Review

2.1 Mind-Body Study

The Mind-Body Study (MBS) was conducted at the Center for Cancer Prevention and Control Research Center at UCLA's Jonsson Comprehensive Cancer Center, with eligible subjects identified through the Los Angeles County Surveillance Epidemiology and End Results (SEER) registry. It was a prospective, longitudinal, observational study designed to elucidate the connection between adjuvant therapies for breast cancer and reports of cognitive dysfunction, commonly reported by women treated with chemotherapy.

Eligible participants were women recently diagnosed with breast cancer, who have undergone primary treatment (surgery, and one of the following: chemotherapy only, radiation only, both chemotherapy and radiation, neither chemotherapy nor radiation) but not endocrine therapy. Details on inclusion/exclusion criteria and study procedures may be found elsewhere (Ganz et al., 2013). The data were collected between 2007 and 2011. Study participants submitted laboratory samples, and took a range of neuropsychological (NP) tests and self-assessment questionnaires, including the PAOFI, at study entry (baseline, denoted T1), 6-month follow-up (denoted T2) and 12-month follow-up (denoted T3). In this dissertation proposal we will focus on analysis of data collected at baseline.

2.2 Use of Patient's Assessment of Own Functioning Inventory (PAOFI) in quality-of-life research

Quality of life (QOL) instruments are designed to assess a person's disposition, attitudes, and concerns regarding various aspects of every-day living. They are usually composed of multiple questions, also called *items*, often organized into subscales or *domains*. Responses to items are commonly measured on an ordinal scale; high correlations between responses in the same domain are expected, and significant intra-domain correlations are common. The complexity and dimensionality of QOL data may be further increased as the data are often collected over multiple time points.

The Patient's Assessment of Own Functioning Inventory (PAOFI) is a self-report QOL instrument designed as a tool for studying the relationship between patients' objective neuropsychological (NP) measures and their self-perceived capabilities and limitations in everyday living activities (Chelune et al., 1986). A modified version of the PAOFI is a questionnaire consisting of 33 items that span four domains of cognitive functioning: memory (items $1-10$; e.g., "How often do you forget events which have occurred in the last day or two?"), language and communication (items $11-19$; e.g., "How often do you have difficulty thinking of names of things?"), sensorimotor abilities (items $20-24$; e.g., "How often do you have difficulty feeling things with your right hand?") and higher-level cognition (items $25-33$; e.g., "How often do you have difficulty finding your way about?"). Each item elicits a response on a 6-point Likert-type scale regarding how often a patient has recently been experiencing a particular type of disability in their everyday life. Specifically, a score of 1 corresponds to "almost always", a score of 2 indicates "very often", a score of 3 reflects "fairly often", a score of 4 represents "once in a while", a score of 5 means "very infrequently", and a score of 6 implies "almost never".

The PAOFI has been used to elicit self-reports of cognitive impairments in a variety of health-research settings; however, the degree of correlation between objective NP mea-

sures and PAOFI varies widely among studies. In breast cancer research, the instrument was reported to correlate with objective neuropsychological tests (Bender et al., 2006, 2008; Pullens et al., 2010; Ganz et al., 2013). The connection with objective test outcomes was not as apparent in a substance-abuse sample (Richardson-Vejlgaard et al., 2009); however, quantity-frequency of drinking was positively correlated with self-perceived cognitive complaints (Shelton and Parsons, 1987). In chronic dialysis patients, researchers found a modest correlation between PAOFI and neuropsychological test scores, but only the subjective PAOFI measure was a significant predictor of Activities of Daily Living scales (Song et al., 2015). Patient's mood disturbances tend to influence self-experience of cognitive impairment. For example, Rourke et al. (1999a,b) found that depression severity explained a large portion of variation in PAOFI scores in HIV patients; Richardson-Vejlgaard et al. (2009) reported that depression accounted for almost half of PAOFI score variance in non-clinical sample. While subjective measures of cognitive functioning might not be perfect surrogates for more costly and time-consuming objective neuropsychological tests, the PAOFI might still provide useful complementary information in certain settings.

It is common to describe findings from the PAOFI using a summation score. Either a global score, the sum of all 33 items, or subscale scores, sums of respective subscale items, are used in analyses. Although the items are scored on a discrete $1-6$ scale, these responses are often dichotomized at a certain cutoff (e.g., scores 1, 2, 3, are assigned a value of 1, and the remaining scores are assigned a value of 0), and the binary result is used to calculate the sum, with a higher PAOFI score being indicative of greater or more extensive cognitive impairment. In the following sections, we review literature on the loss of information due to dichotomization, and the associated potential consequences.

2.3 Dichotomization of non-binary variables

Dichotomization is a process of splitting of a whole into 2 non-overlapping parts. It is a common practice in medical research, where variables originally measured on a continuous or psedo-continuous scale are categorized into binary indicators. Splitting into groups occurs at a prespecified cutoff point, which usually represents a clinically significant value. For example, in a study of the effects of diastolic blood pressure on emergence of headaches, a researcher might group patients into "high BP" and "normal BP" based on some clinically-motivated blood-pressure cutoff, instead of using the full range of diastolic blood pressure values as a predictor. Recently, the journal *Medical Decision Making* has implemented a policy that restricts the use of dichotomization in the journal (Dawson and Weiss, 2012). Despite a great number of publications over the last 70 years urging against dichotomization, it is still routine for dichotomization to be used in several areas of research.

2.4 Information Theory

Research on understanding the consequences of categorizing measurements made on a continuous underlying scale has benefited from insights obtained from the development of information theory.

The field of information theory originated in 1948, when Claude E. Shannon published "A Mathematical Theory of Communication", where he postulated that the amount of information contained in a message is directly proportional to the amount of uncertainty, rather than knowledge, as to what that message conveys (Shannon, 1948). Although the theory was initially developed with applications to electric communication, in the years since it found applications in physics, economics, computer science, linguistics, psychology, statistics, and other disciplines.

Fundamental quantities of information theory are *entropy* and *mutual information*. En-

tropy, or potential information, of a random variable X is defined as the measure of uncertainty regarding its value, and is usually measured in *bits*. One bit is equal to an amount of information necessary to reduce the number of outcomes of X by half. In general, when a discrete random variable X can take on $i = \{1, \ldots, K\}$ mutually exclusive outcomes, each with probability of p_i, the amount of uncertainty associated with each i is equal to

$$H_i = \log_2(1/p_i) \tag{2.1}$$

bits. The entropy of X is defined as a weighted sum of individual uncertainties, i.e.

$$H = \sum_{i=1}^{K} p_i \log_2(1/p_i). \tag{2.2}$$

Shannon further shown that a normally distributed random variable with constant variance stores the largest amount of potential information, estimated as

$$H = \log_2 \sqrt{2\pi e(\frac{\sigma}{\delta})^2}, \tag{2.3}$$

where σ is the standard deviation of X and δ is the precision with which the data were recorded. For example, $\delta = 1$ when height is measured in whole centimeters (e.g., 171 cm); $\delta = 0.1$ when height is measured with precision to the tenth of a centimeter (e.g., 171.3 cm). The more precise is the measurement of X, the more information it has the potential to contain.

Consider a game of "Heads or Tails", and let X denote the outcome of a coin toss. If the coin is fair, there are $i = 2$ possible values of X, each with probability of $p_i = 0.5$, and the amount of uncertainty in a coin flip is $0.5 \log_2 \frac{1}{0.5} + 0.5 \log_2 \frac{1}{0.5} = 1$ bit. Once we know the outcome of a coin toss, the amount of uncertainty is reduced by 1.0 bit and so entropy is equal to 0.0. If a coin is biased and $p_1 \neq p_2$, the amount of uncertainty contained in a single coin flip is less than 1.0 bit, since we possess more information about it than in the equiprobable case.

As the number of possible outcomes of X increases, so does its entropy. Consider drawing a domino from a full set of 36, at random. Because we are drawing at random, getting a

$(1,1)$ is just as likely as getting a $(5,6)$, or a $(6,5)$. So the average uncertainty associated with guessing the outcome of a single draw is $H = \sum_{i=1}^{36} \frac{1}{36} \log_2 36 = \log_2 36 = 5.17$ bits. If we are told that the sum on the chosen domino is even, the number of possible guesses is reduced by half, corresponding to a 1-bit reduction in uncertainty: $\log_2 18 = 4.17$. If we are further assured that at least one half of the domino shows a "3", the number of possibilities is then reduced to a third; hence, the uncertainty is now reduced by more than 1.0 bit and is equal to $\log_2 6 = 2.58$ bits. Once we are told that the sum is equal to "6", we can be certain that our domino is $(3,3)$, and the uncertainty associated with our guess is $\log_2 1 = 0$.

Mutual information is a measure of the amount of information two random variables have in common. Suppose, X and Y are random variables with probability functions $p(x)$ and $p(y)$, respectively. Then, the mutual information of X and Y is defined as

$$I(X;Y) = \sum_x \sum_y p(x,y) \log_2 (\frac{p(x,y)}{p(x)p(y)}), \tag{2.4}$$

where $p(x,y)$ is a joint probability function of X and Y. Because mutual entropy is the amount of information in X that is shared with Y, it can also be written as a difference between marginal entropy of X, $H(X)$, and conditional entropy of X given Y, $H(X|Y)$, i.e.

$$I(X;Y) = H(X) - H(X|Y). \tag{2.5}$$

Consequently, if X and Y are independent, their mutual information is null; when X and Y are collinear, their mutual information is a positive fraction of their respective marginal entropies (Cover and Thomas, 1991).

In the next section, we consider how information theory has contributed to understanding the impact of dichotomizing continuous measures.

2.5 Consequences of dichotomizing continuous variables

Available frameworks for information theory have provided useful insights into studying the consequences of dichotomizing continuous variables in data analysis. In this section, we

review key findings from this literature.

In a study on the overall use of dichotomization in 3 high-quality Psychology journals, MacCallum et al. (2002) found that only about 20% of publications utilizing dichotomization offered some sort of justification for it. Among the explanations offered were: simplification of analyses or presentation of results, and approach used in previous literature. Though enticing in its ability to simplify making a diagnosis, the perceived simplicity of dichotomizing key study outcomes is achieved at a risk of sacrificing other important statistical properties.

Ragland (1992) described the phenomenon of drawing different conclusions from the same data, which occurs because the effect size calculated from a dichotomized outcome is usually not constant over the range of cut-points. Measures of association, such as correlation, were shown to be most affected by "artificial" dichotomization when the split deviated from the 50-50 data allocation (Hunter and Schmidt, 1990). In a study done by Suissa (1991) it was shown that dichotomization of a normally-distributed outcome resulted in loss of efficiency, equivalent to loss in sample size.

Strömberg (1996) conducted a simulation study on the effects of collapsing ordered outcome categories, where he first generated data with 5 ordered categories under the proportional odds model, and then examined odds ratios, Type I errors and powers for each possible collapsed representation of the data. He found that reducing the number of categories led to biased estimates of the effect size and inflated Type I and Type II error rates, with effects being most pronounced when using binary or extreme splits. Taylor et al. (2006) echoed these findings, concluding that the loss of power was particularly dramatic when the collapsed variable was highly skewed, as well as when it had only a few categories.

Beckstead and Beckie (2011) examined information loss that occurs when multiple continuously -measured clinical traits are dichotomized and then aggregated with the purpose of diagnosing metabolic syndrome. Using Shannon's information theory they calculated the amount of information contained in the 5 continuous clinical measures associated with metabolic syndrome, as well the amount of information left in their 5 binary counterparts.

When compared, the amount of information retained in the collapsed measures was only about 12% of the original. As the binary measures were aggregated and the total score was further collapsed to indicate presence or absence of metabolic syndrome, over 98% of collected information was shown to be lost through data transformation. The authors concluded that this common practice of dichotomization-aggregation hinders the progress of medical research, as well as effectiveness of clinical practice.

2.6 Methods for analysis of ordinal quality-of-life data

One way to generalize dichotomous measures is by viewing dichotomous outcomes as collapsed versions of ordinal data. Ordinal data arise in a variety of settings and have been the subject of entire textbooks on data analysis (Agresti, 2010; Johnson and Albert, 2013). In this section, we highlight some key strategies in the literature on analyzing QOL ordinal data.

Johnson and Albert (2013) reviewed latent variable models and item response models, with details on estimation techniques in both frequentist and Bayesian paradigms, as general strategies for analyzing ordinal data. Hays et al. (2000) discussed the advantages of using item response models in QOL research. Other methods of analysis, such as non-model-based analyses, variations on ordinal regression models, and hierarchical ordinal models, were detailed in Agresti (2010).

Ribaudo et al. (2000) employed both a simple model approach and a, more involved, hierarchical modeling strategy in a longitudinal study of self-reports of physical performance in a sample of women undergoing adjuvant treatment for breast cancer. The simple model or "summary statistic" approach consisted of fitting a proportional odds model to subject-specific mean score averaged over all questionnaire items and time points. The multilevel strategy was also done in proportional-odds framework, and included a random effects component to account for repeated measures. They concluded that while the simpler approach

was more common in literature and more familiar to non-statistical audience, the random-effects model was more efficient due to utilizing all of the information in the data.

We now turn to the possibility of analyzing ordinal QOL data using a multi-part model approach, which consists of an initial dichotomization followed by subsequent analysis of the respective parts of the data using appropriate statistical tools.

2.7 Factor analysis for multi-item instruments

Factor analysis is a dimension-reduction technique commonly used in analysis of multi-item instruments designed, such as PAOFI. The goal of factor analysis is to combine instrument items that measure different aspects of an underlying phenomenon into a smaller subset of factors via linear combinations. These factors can then be used as explanatory or response variables in further analyses.

One way to extract factors is by choosing linear combinations of items, X_i's, that maximize the amount of explained variability in X_i's. If there are p items, the linear combination of items that explains the most variability in X_i's is called the first "principal component"; the second principal component is the linear combination that explains the next largest amount of variability yet to be explained; and so on, with the p^{th} principal component consisting of the remaining linear combination of items orthogonal to the first *(p-1)* principal components. A measure of relative proportion of variability explained by a factor is called an eigenvalue. While the number of possible principal components is equal to the number of X_i's, only the first few are potentially informative, due to the nature of the extraction process. When choosing the number of principal components to retain for further analyses, a common strategy is to keep only those combinations whose eigenvalues are greater than 1; however, models with fewer or more principal components may be of interest. Other strategies consider the proportion of variance explained by the ensemble of factors, gaps between successive eigenvalues (signaling a substantial drop in the amount of variance explained by

the new factor), and/or the interpretability of the ensemble of factors on subjective grounds.

Because there is not a unique solution to identifying factors in factor analysis since any rotation of a set of factors will explain the same amount of variability in the data, a typical next step is to transform the chosen linear combinations into interpretable factors by using a factor rotation. A common method that ensures that resultant factors are uncorrelated is called "varimax rotation". This method favors factor loadings, coefficients of original variables, that are close to either 1 or -1, which is indicative of the strength of the relationship between the instrument item and the corresponding construct it aims to measure, or to 0, suggestive of the item not having a strong relationship with the given construct (Afifi et al., 2011).

There are other strategies for extracting factors from multivariate data that can be viewed as alternatives to principal-component extraction as well as other strategies for rotation that can be viewed as alternatives to varimax rotation. Although different approaches can give rise to differently interpreted factors, it is also well accepted that there is not a single "right" way to perform factor analysis, and the combination of principal-component analysis with varimax rotation is a widely-used strategy for developing composite variables from multidimensional instruments. In Chapter 4, we will use principal-component analysis accompanied by varimax rotation to explore different possible approaches for translating the 33 PAOFI items into a smaller number of interpretable constructs.

2.8 Multipart models in health research settings

Another strategy for analyzing ordinal categorical data that could be appropriate in certain contexts is to analyze outcomes in a multi-part model. This might be especially appropriate if certain levels can be assumed to depend on a dichotomization in a hierarchical framework, as with multi-part models in health-services research, where some proportion of people might have no service utilization during a given period, while among those with any service uti-

lization there would be a distribution of the amount, and the predictors of the respective parts might differ (Duan et al., 1984). Another example is the idea behind "cure models" in survival analysis, where some portion of people are effectively cured by a treatment, the complementary proportion could be modeled within a traditional survival-analysis framework and the predictors for the respective parts of the model could differ (Sy and Taylor, 2000).

A two-part model was developed as a tool for analysis of limited dependent variables, which are characterized by distributions with significant probability mass located at the lower bound, usually zero. Because the zeros and the non-zero values are typically generated by different mechanisms, separate models allowing for different predictors are used to fit the data. First, the probability of observing a non-zero outcome is modeled as

$$\phi(y > 0) = P(y > 0|\boldsymbol{x_1}) = F(\boldsymbol{x_1}\boldsymbol{\beta}), \tag{2.6}$$

where $\boldsymbol{x_1}$ is a vector of covariates, $\boldsymbol{\beta}$ is a vector of parameter estimates, and F is the link function, (e.g., logit). Second part of the model is fit to the non-zero values as

$$\phi(y|y > 0) = g(\boldsymbol{x_2}\boldsymbol{\gamma}), \tag{2.7}$$

where $\boldsymbol{x_2}$ is a vector of covariates, $\boldsymbol{\gamma}$ is a vector of parameter estimates, and g is a density function (e.g., Gaussian).

Stata function, twopm, is available for fitting of two-part models. In addition to producing standard parameter estimates and inference, it has the ability to perform joint statistical tests of parameters from the two parts of the model, as well as calculate predicted values and the associated standard errors for the entire sample (Belotti et al., 2015).

In Chapter 5 we develop notation for an approach that could be adapted to analyze PAOFI data.

CHAPTER 3

Preliminary Analyses

The analyses of the PAOFI in Ganz et al. (2013) for analyzing data from the previously described Mind-Body Study (MBS) invoked the strategy that has been widely used elsewhere in the neuropsychology literature, where values of 1,2,3 are recoded as 1 and 4,5,6 are recoded as 0. Based on information theory and related findings in the statistical literature, there are grounds to believe that additional insights would be available by using the available data differently.

In this chapter, we summarize some findings from preliminary analyses in this direction. As a first step, we explore both alternative cut-points and the possibility of treating the ordinal PAOFI measures as if they were continuous. Specifically, we present summary statistics, Spearman's rank correlation coefficients, and histograms of individual PAOFI items; histograms of composite variables based on dichotomized and undichotomized PAOFI items, and Pearson correlations between them; results of an investigation using Shannon's information theory; results of regressing composite variables based on dichotomized and undichotomized PAOFI items on treatment categories (chemotherapy only, radiation only, both chemotherapy and radiation, neither chemotherapy nor radiation).

3.1 Summary statistics and graphical representation of PAOFI items and corresponding domain scales

Histograms of the 33 PAOFI items split into 4 domains (memory, language and communication, sensorimotor, higher-level cognition) are presented in Figures 3.1-3.4. Most items distributions are skewed right, indicating that majority of patients did not experience any particular problem often, if at all. This is corroborated by summary statistics (Tables 3.1-3.4), as the median score on all items is 4 ("once in a while") or higher ("very infrequently", "almost never"). As expected, Spearman's rank correlation coefficients, ranging between 0.19 and 0.78, indicate a degree of collinearity between items on the same scale, with the highest correlations (> 0.7) observed between 4 items in the higher-level cognition domain of the PAOFI. Most items were scored by all 189 subjects on the study.

In the remaining figures and tables the following notation is used:

$W^{(1)}$ indicates sum of dichotomized items, where items with score 1 were assigned a score of 1, and items with scores 2,3,4,5,6 were assigned a score of 0;

$W^{(2)}$ indicates sum of dichotomized items, where items with scores 1,2 were assigned a score of 1, and items with scores 3,4,5,6 were assigned a score of 0;

$W^{(3)}$ indicates sum of dichotomized items, where items with scores 1,2,3 were assigned a score of 1, and items with scores 4,5,6 were assigned a score of 0;

$W^{(4)}$ indicates sum of dichotomized items, where items with scores 1,2,3,4 were assigned a score of 1, and items with scores 5,6 were assigned a score of 0;

$W^{(5)}$ indicates sum of dichotomized items, where items with scores 1,2,3,4,5 were assigned a score of 1, and items with score 6 were assigned a score of 0;

Q indicates sum of reverse-coded items on the original 6-point scale.

Reverse coding is used to maintain positive association between scales composed of variously dichotomized and undichotomized items.

The choice of cut-points for dichotomization of scale items can have a dramatic effect on the resulting composite variable. Tables 3.5 and 3.6 show Pearson correlation coefficients between composite variables based on uncollapsed item scores (Q), with rows corresponding to sums of the given number of ordinal items on a scale from 1 to 6, and sums of dichotomized items under different dichotomization rules ($W^{(1)} - W^{(5)}$). The correlations tend to be lower for dichotomizing between 1 and 2-6 than dichotomizing between 1-5 and 6, reflecting that participants often don't utilize the range of possible values and that 1("almost always") is not a particularly common item score. The highest correlations tend to be for dichotomizations based on a split of 1-4 vs 5-6. The pattern is similar for items on the original ordinal scale (Table 3.5) and for log transformations of the sums of the 6-point ordinal items (Table 3.6). Figures 3.5-3.8 show distributions of subscale sums formed by dichotomized ($W^{(1)} - W^{(5)}$) and undichotomized (Q) items. Because of high degree of skewness in the original item distributions, composite sums exhibit skewness as well, with $W^{(1)} - W^{(3)}$ not spanning the whole range of possible values on all scales but higher-level cognition.

In passing, we note that dichotomization in some settings is based on the median split of the data. However, for some PAOFI items this would imply using $W^{(4)}$ or $W^{(5)}$, rather than the more common $W^{(3)}$, and for other PAOFI items it would require addressing the ambiguity of what to do when the median score was 6. This underscores that in settings with multi-domain instruments using a dichotomization rule based on median splits may be a source of confusion as well as difficulty in interpretation.

3.2 Information loss due to collapsing non-binary outcomes

In this section we apply information theory introduced in Section 2.4, to estimate the amount of potential information loss in the MBS data when PAOFI items are dichotomized prior to

being aggregated into respective subscale scores. The uncertainties in Memory subscale responses are calculated before and after the responses are dichotomized (scores 1,2,3 assigned a score of 1, scores 4,5,6 assigned a score of 0); mutual entropy is estimated to compare potential information contained in original vs dichotomized responses. These calculations do not take into account correlations between items; however, the effects of correlations are not expected to change these results dramatically.

The results, presented in Table 3.7, show the amount of uncertainty in each Memory item, as well as the total amount of potential information, H, in the subscale before and after dichotomization, assuming no correlation between items. Calculations are done using equation 2.2, where p_i is the proportion of respondents endorsing a particular response. Because original data were recorded on a discrete scale from 1 to 6, the entropy in originally-scaled items, $H_{original} = 19.346$, is obtained by summing over 60 (10 questions with 6 responses each) uncertainty components (not shown). Analogously, $H_{dichotomized} = 5.676$ is calculated using information from 20 possible contributions. In addition, because the underlying scale for each response is continuous, calculations under the assumption of normality are performed to obtain a benchmark for the maximum information possible. Sample standard deviation of each item and precision of $\delta = 1$ are used in equation 2.3. Under the most favorable condition, if each item were normally distributed, $H_{normal} = 29.619$.

To estimate the amount of information lost due to data transformation, mutual information, $I(X;Y)$, is examined. Let X represent a random variable from the original distribution, and Y - a random variable representing data after dichotomization. Because Y is completely determined by X, $I(X;Y) = H(Y)–H(Y|X) = H(Y)$. So, the amount of information knowing Y provides about X, and vice versa, is 5.676 bits. This corresponds to a loss of 19.346–5.676 = 13.67 bits or over 70% of collected information.

The data suggest that only about 30% of the information in Memory scale items is retained after individual items are dichotomized and before they are aggregated into respective domain scores. While information theory does not take into account the quality of informa-

tion, discarding 70% of data collected from a well-designed instrument is equivalent to losing over 2/3 of the sample size. It is, therefore, recommended that responses to the PAOFI are not dichotomized prior to conducting statistical analyses.

3.3 Multiple linear regression of composite outcomes on predictors

To illustrate key ideas, we follow the analysis approaches used in Ganz et al. (2013) paper, focusing on whether the relationships between variables are revealed as significant in OLS regression and on the magnitudes of implied effects. We present the results of linear regression of log-transformed PAOFI subscale sums composed of dichotomized ($W^{(1)} - W^{(5)}$) and undichotomized (Q) variables on treatment options (chemotherapy only, radiation only, both chemotherapy and radiation, neither chemotherapy nor radiation) in Tables 3.8-3.12.

Table 3.8 illustrates that the choice of cut-point has an impact on the significance and magnitude of the estimated effects. While results are mostly similar among models $W^{(2)} - W^{(4)}$, in comparison, models $W^{(1)}$ and $W^{(5)}$ exhibit smaller effects and increased p-values. Because effect sizes and standard errors are not directly comparable between models $W^{(1)} - W^{(5)}$ and Q, Table 3.9 shows estimates and standard errors when all outcomes are standardized. As expected, significance of the results is not affected by this transformation, but estimates of the coefficients and standard errors of model Q are similar to those from models $W^{(2)} - W^{(5)}$. In the Language and Communication domain, results are most similar between models Q and $W^{(4)}$; in Higher-Level Cognition domain models $W^{(3)}$ and $W^{(5)}$ exhibit significance patterns similar to those in model Q.

Most models agree that there is a possible indication of increased memory complaints when comparing chemotherapy-only with neither-chemo-nor-radiation patients ($0.05 < p < 0.1$), and there is a strong indication of more complaints in the chemo-and-radiation group when compared to neither-chemo-nor-radiation patients ($p < 0.01$). Patients who have only

undergone radiation tend to have significantly fewer memory complaints than do patients who have undergone both chemo and radiation ($p < 0.01$). Overall, patients seem to have fewer complaints regarding memory, language and communication, and higher-level cognition, if they were exposed to radiation treatment rather than chemotherapy only or both chemo and radiation; patients on chemo only or on both chemo and radiation tend to report significantly more complaints than patients who received neither therapy. These results are not directly comparable to those in Ganz et al. (2013) as our analyses are not adjusted for covariates.

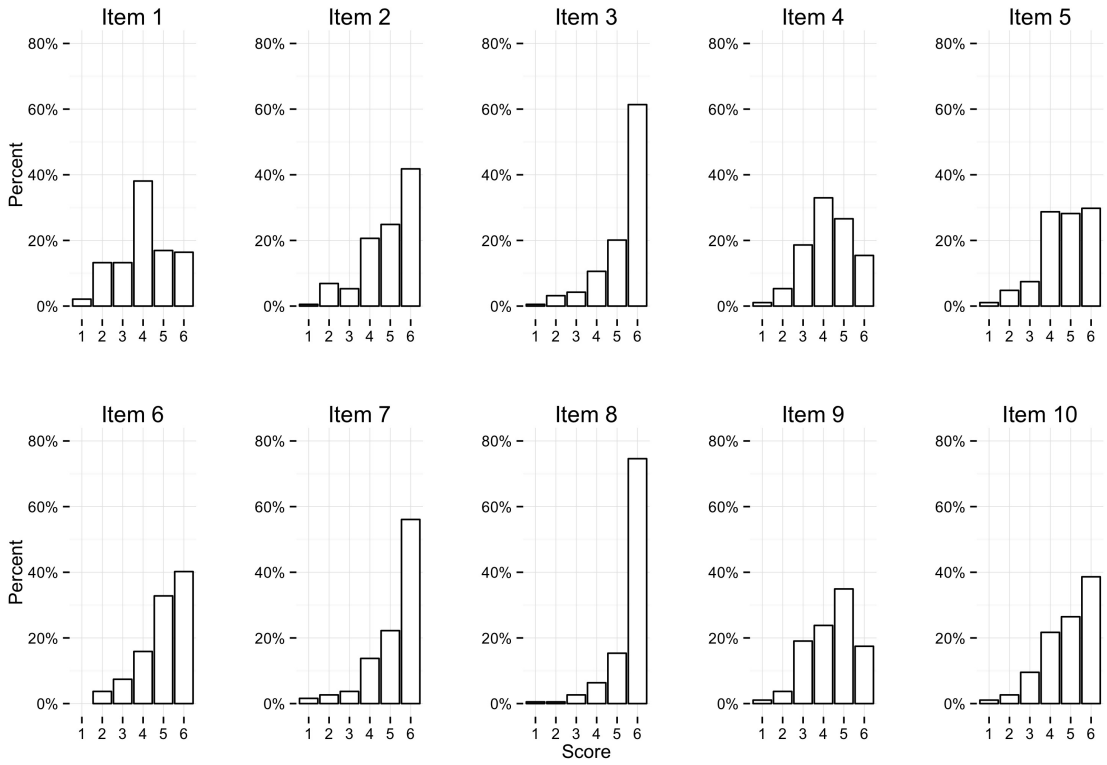

Figure 3.1: Distributions of items from the Memory domain.

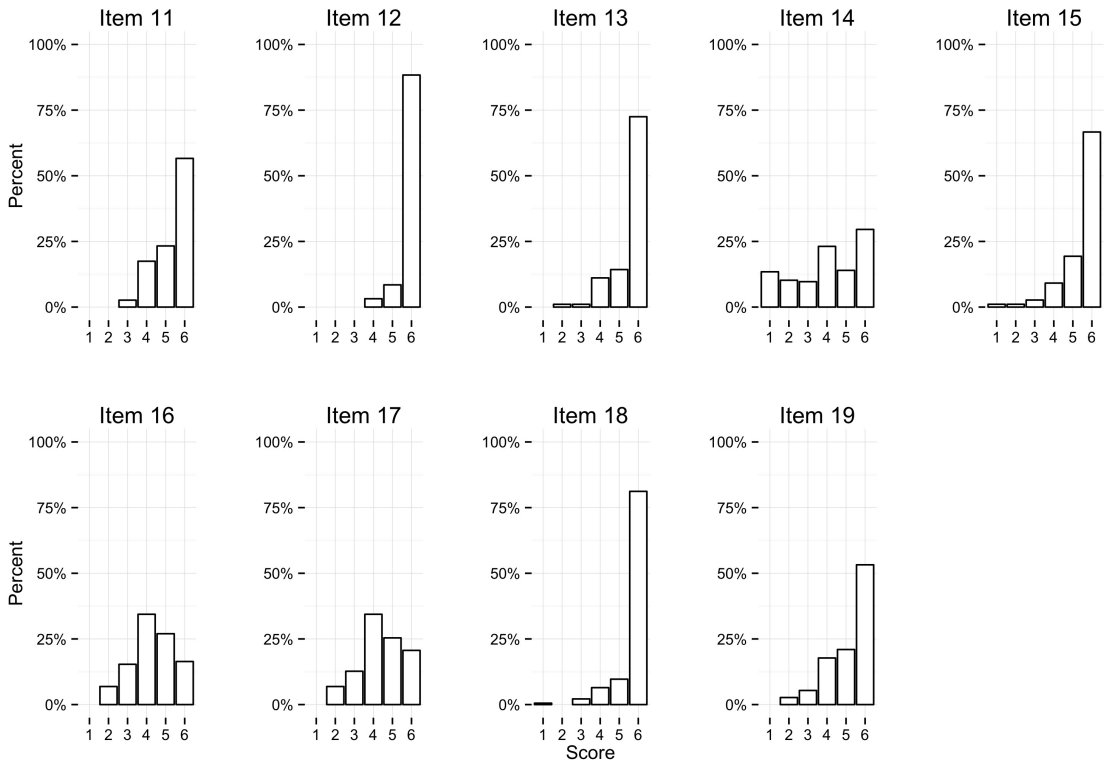

Figure 3.2: Distributions of items from the Language & Communication domain.

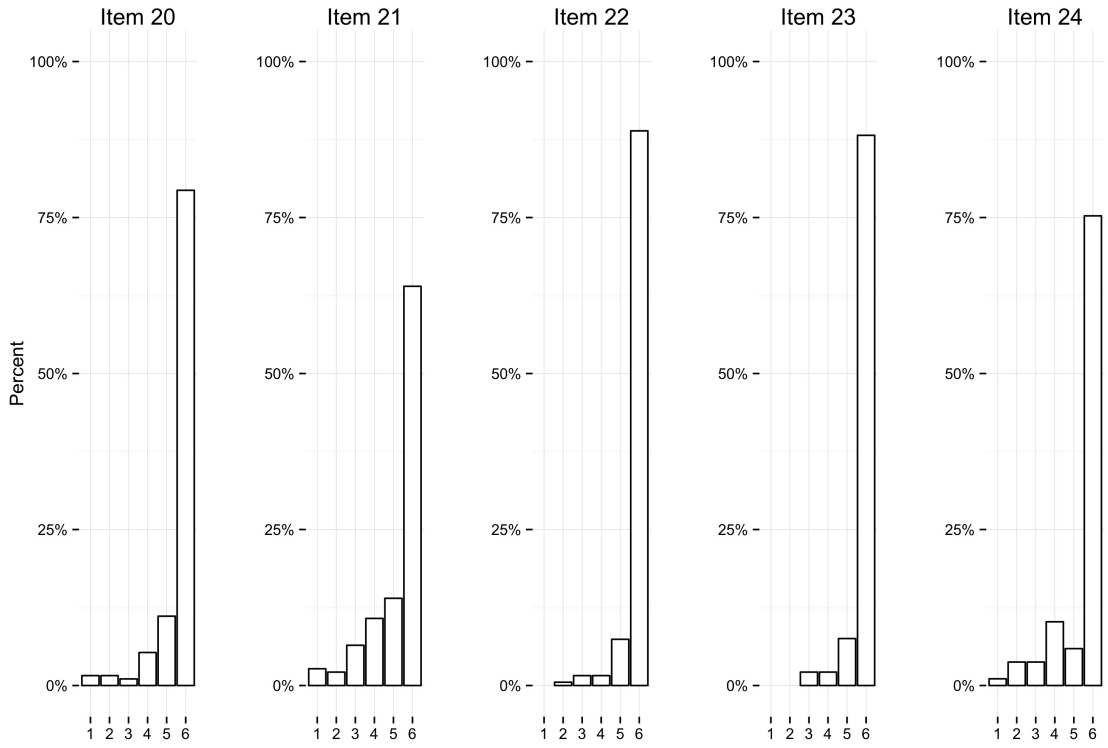

Figure 3.3: Distributions of items from the Sensorimotor domain.

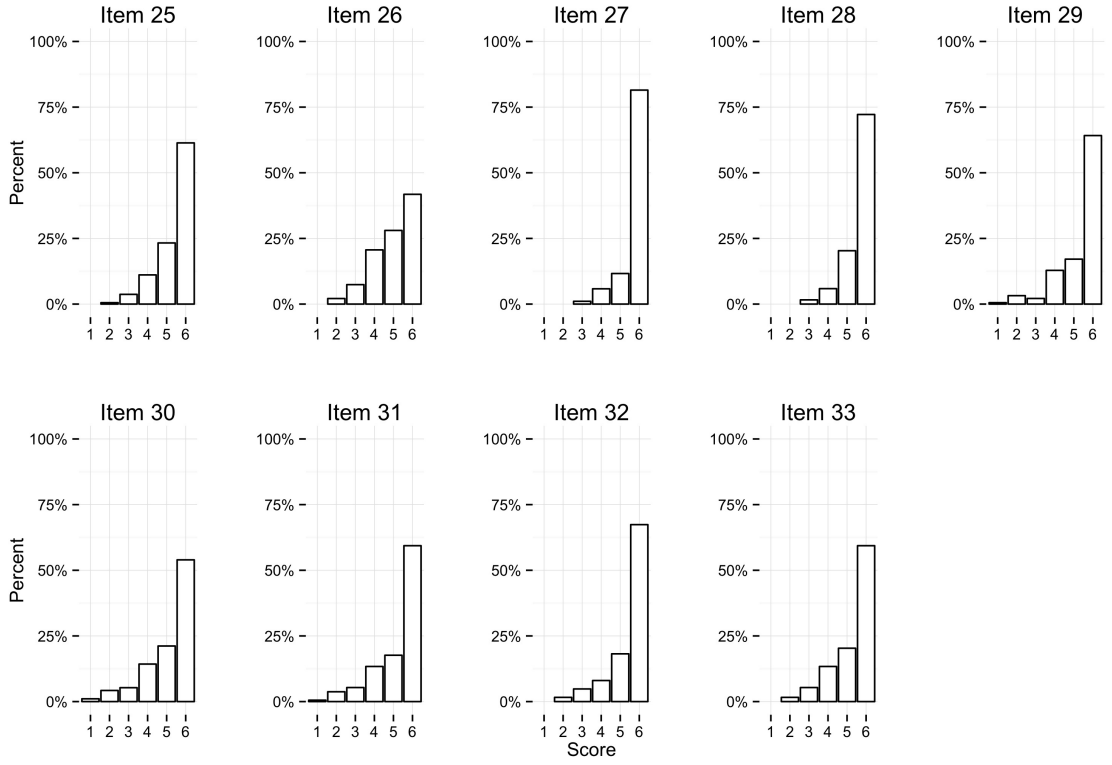

Figure 3.4: Distributions of items from the Higher-Level Cognition domain.

Table 3.1: Spearman's rank correlation coefficients and summary statistics for items comprising PAOFI Memory domain.

PAOFI	Spearman's Rank Correlation Coefficient, r_s										Summary Statistics			
	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	n	median	$\bar{x}$	s
Item 1	1.00										189	4	4.0	1.30
Item 2	0.66	1.00									189	5	4.9	1.24
Item 3	0.51	0.60	1.00								188	6	5.3	1.08
Item 4	0.50	0.50	0.43	1.00							188	4	4.2	1.15
Item 5	0.30	0.33	0.50	0.60	1.00						189	5	4.7	1.18
Item 6	0.38	0.35	0.32	0.34	0.24	1.00					189	5	5.0	1.09
Item 7	0.39	0.44	0.35	0.32	0.22	0.52	1.00				189	6	5.2	1.14
Item 8	0.36	0.32	0.30	0.41	0.26	0.34	0.38	1.00			189	6	5.6	0.84
Item 9	0.38	0.34	0.31	0.42	0.19	0.37	0.37	0.45	1.00		189	5	4.4	1.15
Item 10	0.39	0.42	0.33	0.40	0.25	0.44	0.49	0.40	0.55	1.00	189	5	4.9	1.17

Table 3.2: Spearman's rank correlation coefficients and summary statistics for items comprising PAOFI Language & Communication domain.

PAOFI	Spearman's Rank Correlation Coefficient, r_s									Summary Statistics			
	Item 11	Item 12	Item 13	Item 14	Item 15	Item 16	Item 17	Item 18	Item 19	n	median	$\bar{x}$	s
Item 11	1.00									189	6	5.3	0.86
Item 12	0.41	1.00								189	6	5.8	0.44
Item 13	0.61	0.48	1.00							189	6	5.6	0.81
Item 14	0.32	0.27	0.25	1.00						187	4	4.0	1.76
Item 15	0.35	0.35	0.24	0.25	1.00					189	6	5.5	0.96
Item 16	0.32	0.23	0.20	0.30	0.27	1.00				189	4	4.3	1.13
Item 17	0.27	0.19	0.26	0.30	0.30	0.70	1.00			189	4	4.4	1.15
Item 18	0.21	0.25	0.14	0.25	0.28	0.27	0.31	1.00		188	6	5.7	0.77
Item 19	0.39	0.30	0.28	0.31	0.24	0.37	0.41	0.46	1.00	188	6	5.2	1.07

Table 3.3: Spearman's rank correlation coefficients and summary statistics for items comprising PAOFI Sensorimotor domain.

PAOFI	Spearman's Rank Correlation Coefficient, r_s					Summary Statistics			
	Item 20	Item 21	Item 22	Item 23	Item 24	n	median	$\bar{x}$	s
Item 20	1.00					189	6	5.6	0.96
Item 21	0.30	1.00				188	6	5.2	1.25
Item 22	0.35	0.21	1.00			189	6	5.8	0.58
Item 23	0.19	0.40	0.66	1.00		188	6	5.8	0.57
Item 24	0.19	0.22	0.31	0.30	1.00	188	6	5.4	1.16

Table 3.4: Spearman's rank correlation coefficients and summary statistics for items comprising PAOFI Higher-Level Cognition domain.

PAOFI	Spearman's Rank Correlation Coefficient, r_s									Summary Statistics			
	Item 25	Item 26	Item 27	Item 28	Item 29	Item 30	Item 31	Item 32	Item 33	n	median	$\bar{x}$	s
Item 25	1.00									189	6	5.4	0.87
Item 26	0.62	1.00								189	5	5.0	1.06
Item 27	0.52	0.41	1.00							189	6	5.7	0.61
Item 28	0.40	0.30	0.61	1.00						189	6	5.6	0.67
Item 29	0.54	0.41	0.45	0.42	1.00					188	6	5.3	1.06
Item 30	0.56	0.53	0.41	0.42	0.62	1.00				189	6	5.1	1.20
Item 31	0.59	0.54	0.47	0.47	0.61	0.78	1.00			188	6	5.2	1.16
Item 32	0.50	0.39	0.53	0.62	0.54	0.58	0.66	1.00		189	6	5.4	0.94
Item 33	0.62	0.49	0.46	0.46	0.57	0.64	0.74	0.73	1.00	188	5	5.3	1.00

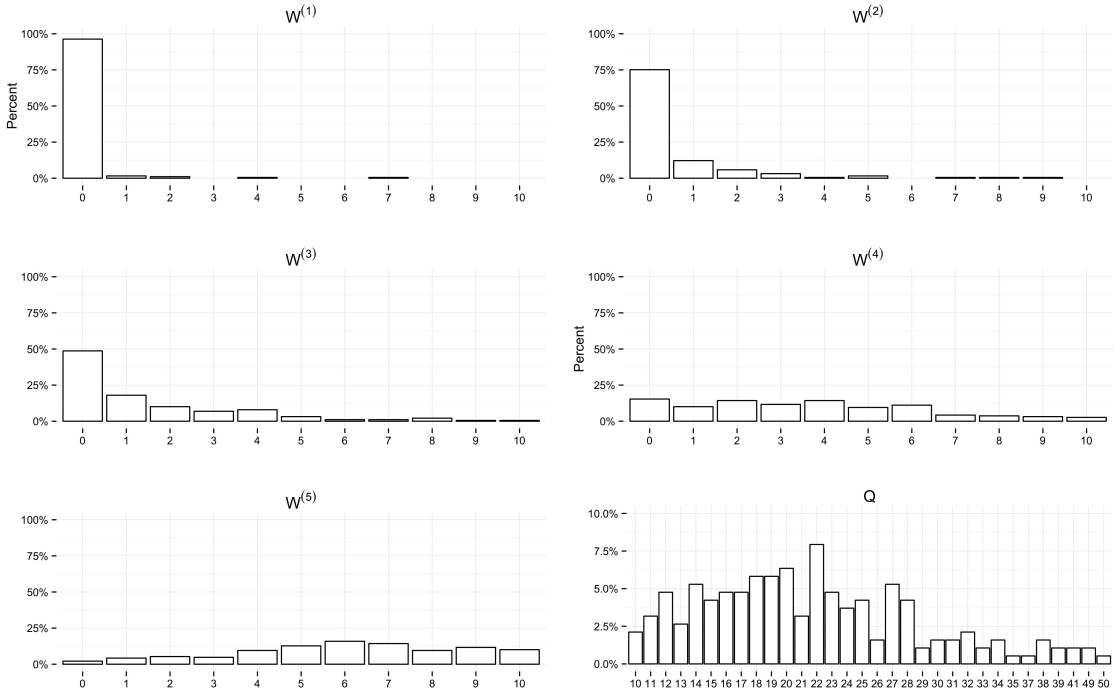

Figure 3.5: Distributions of composite variables based on dichotomized ($W^{(1)} - W^{(5)}$) and undichotomized (Q) items from the Memory domain.

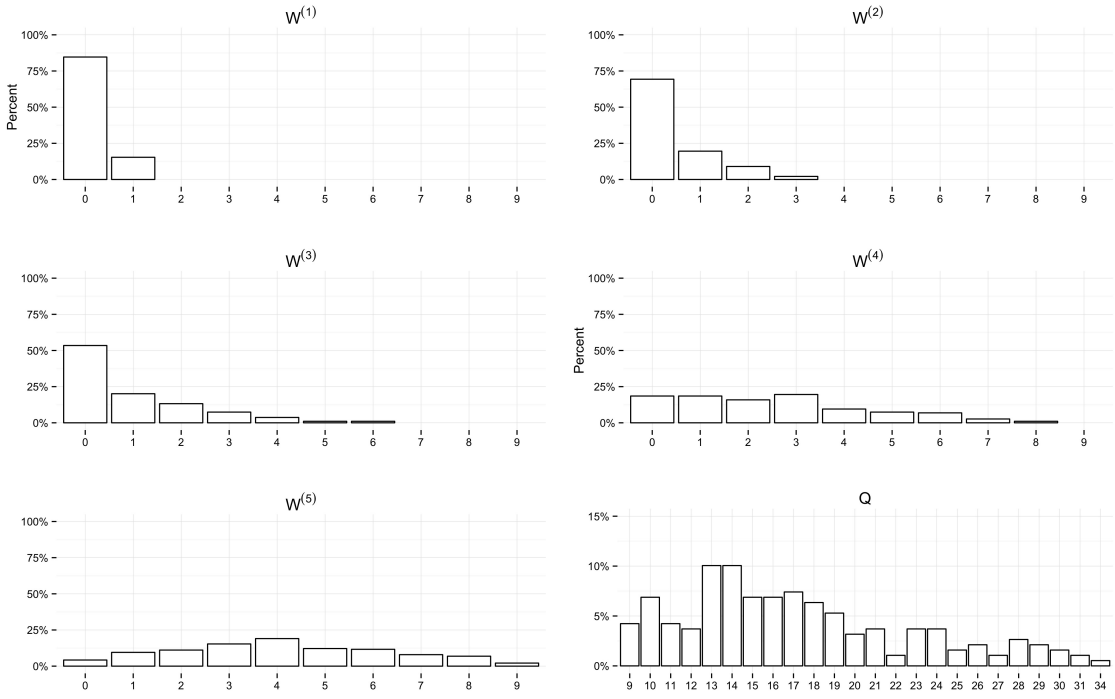

Figure 3.6: Distributions of composite variables based on dichotomized ($W^{(1)} - W^{(5)}$) and undichotomized (Q) items from the Language & Communication domain.

Table 3.5: Pearson correlation coefficients relating PAOFI subscale sums composed of items on the original 6-point scale with sums composed of variously dichotomized items.

PAOFI Subscale	Dichotomization Rule				
	$W^{(1)}$	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$	$W^{(5)}$
Memory, 10 items	0.423	0.728	0.895	0.933	0.845
Language & Communication, 9 items	0.500	0.716	0.833	0.933	0.885
Sensorimotor, 5 items	0.559	0.724	0.874	0.909	0.887
Higher-Level Cognition, 9 items	0.396	0.574	0.824	0.945	0.904

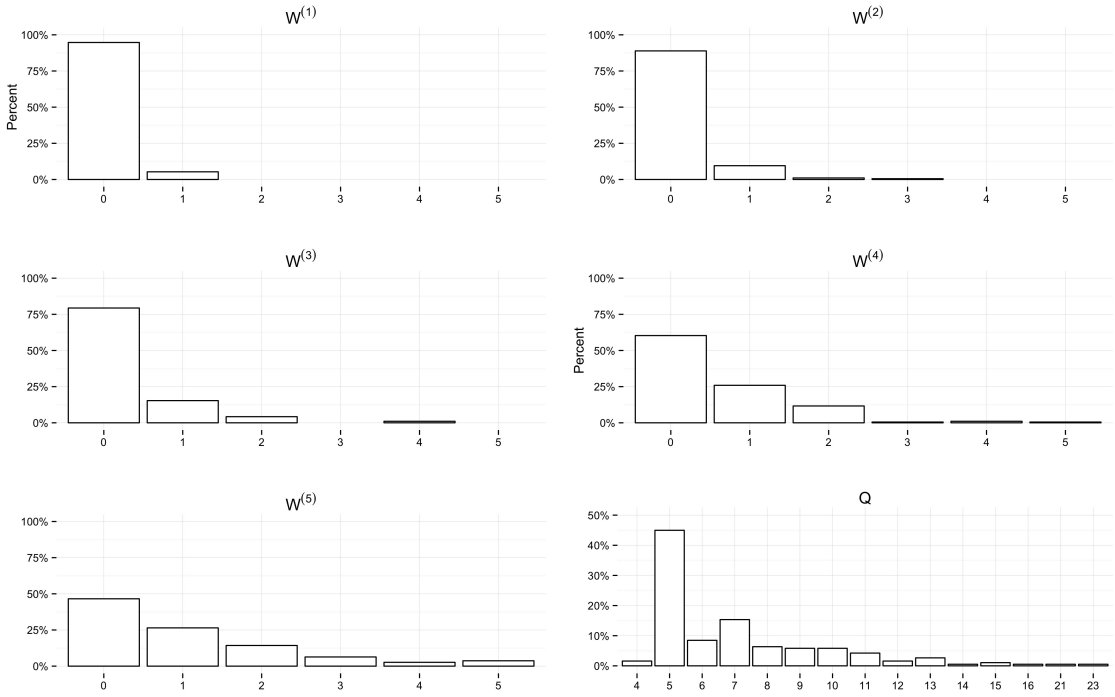

Figure 3.7: Distributions of composite variables based on dichotomized ($W^{(1)} - W^{(5)}$) and undichotomized (Q) items from the Sensorimotor domain.

Table 3.6: Pearson correlation coefficients relating log-transformed PAOFI subscale sums composed of items on the original 6-point scale with log-transformed sums composed of variously dichotomized items.

PAOFI Subscale	Dichotomization Rule				
	$W^{(1)}$	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$	$W^{(5)}$
Memory, 10 items	0.347	0.668	0.839	0.924	0.861
Language & Communication, 9 items	0.472	0.670	0.785	0.915	0.879
Sensorimotor, 5 items	0.488	0.638	0.815	0.901	0.920
Higher-Level Cognition, 9 items	0.326	0.543	0.779	0.937	0.907

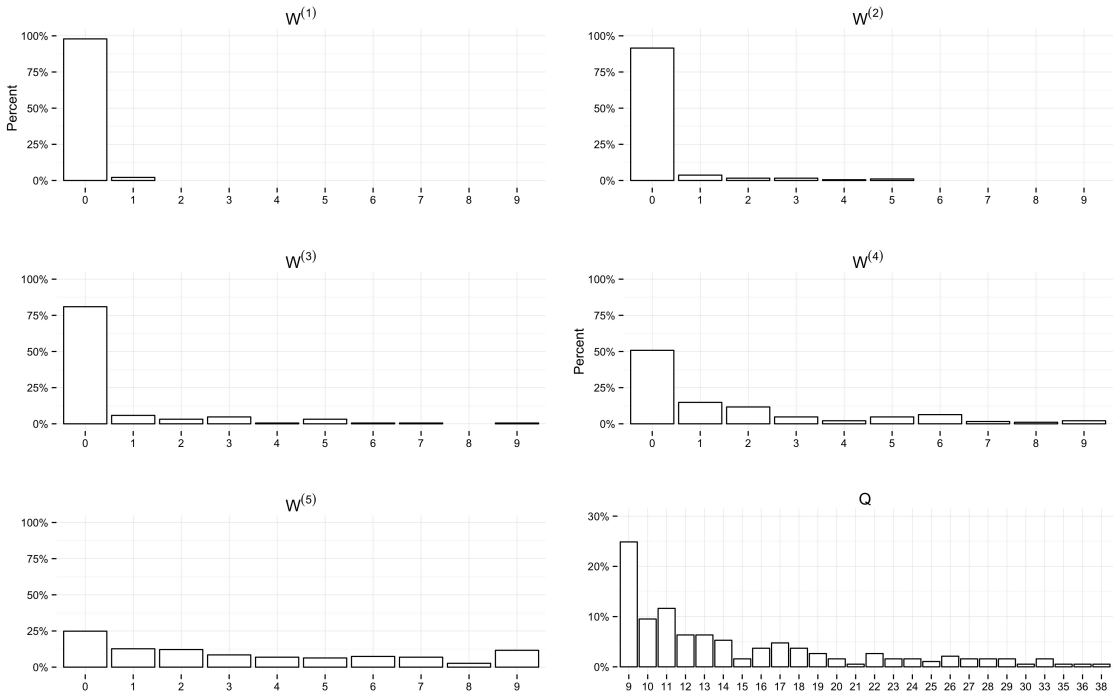

Figure 3.8: Distributions of composite variables based on dichotomized ($W^{(1)} - W^{(5)}$) and undichotomized (Q) items from the Higher-Level Cognition domain.

Table 3.7: Entropy in dichotomized and originally-scaled items from the PAOFI Memory subscale, assuming no correlation between items ($n = 188$).

Item	Dichotomized		Original, 6-point	Normal	
	$P(x = 0)$	$H_{dichotomized}$, bits	$H_{original}$, bits	s	H_{normal}, bits
1	0.714	0.863	2.282	1.298	3.198
2	0.873	0.549	2.025	1.238	3.112
3	0.921	0.400	1.631	1.082	2.887
4	0.750	0.811	2.198	1.150	2.984
5	0.867	0.566	2.111	1.177	3.024
6	0.889	0.503	1.932	1.094	2.904
7	0.921	0.400	1.754	1.141	2.972
8	0.963	0.228	1.202	0.837	2.533
9	0.762	0.792	2.164	1.152	2.987
10	0.868	0.564	2.047	1.174	3.019
Total:		5.676	19.346		29.619

Table 3.8: Effect size and standard errors from regressions of log-transformed Memory scores on treatment ($n = 188$).

Contrast	Regressions of log(1+Memory score)					
	$W^{(1)}$	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$	$W^{(5)}$	Q
Radiation *vs*	0.01	0.02	0.02	0.17	-0.02	0.04
No Chemotherapy & No Radiation	(0.053)	(0.111)	(0.152)	(0.157)	(0.113)	(0.076)
Chemotherapy *vs*	0.10	***0.31*****	***0.35****	***0.36****	0.05	***0.17****
No Chemotherapy & No Radiation	(0.069)	***(0.144)***	***(0.196)***	***(0.203)***	(0.146)	***(0.099)***
Chemotherapy & Radiation *vs*	0.07	***0.33******	***0.52******	***0.50******	0.17	***0.23******
No Chemotherapy & No Radiation	(0.052)	***(0.109)***	***(0.148)***	***(0.153)***	(0.110)	***(0.074)***
Radiation *vs*	-0.09	***-0.29*****	***-0.33****	-0.19	-0.06	-0.13
Chemotherapy	(0.060)	***(0.126)***	***(0.172)***	(0.178)	(0.128)	(0.086)
Chemotherapy *vs*	0.04	-0.02	-0.17	-0.194	-0.12	-0.06
Chemotherapy & Radiation	(0.059)	(0.123)	(0.168)	(0.174)	(0.125)	(0.085)
Radiation *vs*	-0.06	***-0.31******	***-0.50******	***-0.33******	***-0.18*****	***-0.19******
Chemotherapy & Radiation	(0.040)	***(0.083)***	***(0.113)***	***(0.117)***	***(0.084)***	***(0.057)***

Note: Bold font indicates significance; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 3.9: Effect size and standard errors from regressions of log-transformed standardized Memory scores on treatment ($n = 188$).

Contrast	Regressions of log(1+ standardized Memory score)					
	$W^{(1)}$	$W^{(2)}$	$W^{3)}$	$W^{(4)}$	$W^{(5)}$	Q
Radiation *vs*	0.05	0.04	0.03	0.24	-0.03	0.12
No Chemotherapy & No Radiation	(0.226)	(0.217)	(0.215)	(0.220)	(0.225)	(0.219)
Chemotherapy *vs*	0.44	***0.60*****	***0.50****	***0.51****	0.09	***0.49****
No Chemotherapy & No Radiation	(0.292)	***(0.281)***	***(0.277)***	***(0.285)***	(0.291)	***(0.283)***
Chemotherapy & Radiation *vs*	0.29	***0.65******	***0.73******	***0.70******	0.33	***0.66******
No Chemotherapy & No Radiation	(0.220)	***(0.212)***	***(0.209)***	***(0.215)***	(0.219)	***(0.214)***
Radiation *vs*	-0.40	***-0.56*****	***-0.47****	-0.27	-0.12	-0.38
Chemotherapy	(0.255)	***(0.246)***	***(0.243)***	(0.249)	(0.255)	(0.248)
Chemotherapy *vs*	0.16	-0.04	-0.23	-0.19	-0.24	-0.17
Chemotherapy & Radiation	(0.250)	(0.241)	(0.238)	(0.244)	(0.249)	(0.243)
Radiation *vs*	-0.24	***-0.61******	***-0.71******	***-0.46******	***-0.36*****	***-0.55******
Chemotherapy & Radiation	(0.169)	***(0.162)***	***(0.160)***	***(0.164)***	***(0.168)***	***(0.164)***

Note: Bold font indicates significance; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 3.10: Effect size and standard errors from regressions of log-transformed Language & Communication scores on treatment ($n = 182$).

Contrast	Regressions of log(1+L&C score)					
	$W^{(1)}$	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$	$W^{(5)}$	Q
Radiation vs	-0.06	-0.08	-0.06	0.04	0.07	0.00
No Chemotherapy & No Radiation	(0.057)	(0.094)	(0.130)	(0.143)	(0.118)	(0.071)
Chemotherapy vs	-0.02	0.04	0.08	***0.36****	***0.29****	***0.16****
No Chemotherapy & No Radiation	(0.073)	(0.122)	(0.168)	***(0.184)***	***(0.153)***	***(0.092)***
Chemotherapy & Radiation vs	0.01	0.06	0.15	***0.37*****	***0.34*****	***0.19*****
No Chemotherapy & No Radiation	(0.055)	(0.092)	(0.127)	***(0.139)***	***(0.115)***	***(0.069)***
Radiation vs	-0.04	-0.12	-0.14	***-0.33****	-0.22	***-0.16****
Chemotherapy	(0.064)	(0.107)	(0.147)	***(0.161)***	(0.134)	***(0.080)***
Chemotherapy vs	-0.03	-0.02	-0.7	-0.01	-0.05	-0.03
Chemotherapy & Radiation	(0.063)	(0.105)	(0.144)	(0.158)	(0.131)	(0.079)
Radiation vs	-0.07	***-0.14****	***-0.21****	***-0.33*****	***-0.26*****	***-0.18*****
Chemotherapy & Radiation	(0.042)	***(0.070)***	***(0.097)***	***(0.106)***	***(0.088)***	***(0.053)***

Note: Bold font indicates significance; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 3.11: Effect size and standard errors from regressions of log-transformed Sensorimotor scores on treatment ($n = 188$).

Contrast	Regressions of log(1+Sensorimotor score)					
	$W^{(1)}$	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$	$W^{(5)}$	Q
Radiation vs	-0.02	-0.01	-0.07	-0.10	***-0.26***	-0.12
No Chemotherapy & No Radiation	(0.035)	(0.057)	(0.080)	(0.101)	***(0.127)***	(0.077)
Chemotherapy vs	-0.03	-0.01	-0.03	0.16	-0.08	-0.01
No Chemotherapy & No Radiation	(0.046)	(0.073)	(0.103)	(0.130)	(0.164)	(0.099)
Chemotherapy & Radiation vs	0.02	0.04	0.01	0.08	-0.04	0.02
No Chemotherapy & No Radiation	(0.034)	(0.055)	(0.078)	(0.098)	(0.124)	(0.075)
Radiation vs	0.04	-0.00	-0.04	***-0.26***	-0.18	-0.12
Chemotherapy	(0.040)	(0.064)	(0.090)	***(0.114)***	(0.144)	(0.087)
Chemotherapy vs	-0.05	-0.04	-0.05	0.07	-0.04	-0.02
Chemotherapy & Radiation	(0.039)	(0.063)	(0.089)	(0.112)	(0.141)	(0.085)
Radiation vs	-0.00	-0.05	-0.08	***-0.18***	***-0.22***	***-0.14***
Chemotherapy & Radiation	(0.026)	(0.042)	(0.060)	***(0.075)***	***(0.095)***	***(0.057)***

Note: Bold font indicates significance; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 3.12: Effect size and standard errors from regressions of log-transformed Higher-Level Cognition scores on treatment ($n = 182$).

Contrast	Regressions of log(1+HLC score)					
	$W^{(1)}$	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$	$W^{(5)}$	Q
Radiation vs	0.01	0.07	0.08	-0.01	0.02	0.03
No Chemotherapy & No Radiation	(0.023)	(0.073)	(0.119)	(0.166)	(0.182)	(0.086)
Chemotherapy vs	0.04	***0.27******	***0.33****	0.31	***0.54****	***0.27****
No Chemotherapy & No Radiation	(0.029)	***(0.094)***	***(0.154)***	(0.215)	***(0.235)***	***(0.111)***
Chemotherapy & Radiation vs	0.02	0.10	***0.31*****	***0.41****	***0.49*****	***0.26*****
No Chemotherapy & No Radiation	(0.022)	(0.071)	***(0.116)***	***(0.162)***	***(0.177)***	***(0.084)***
Radiation vs	-0.02	***-0.21****	***-0.25****	***-0.32****	***-0.52****	***-0.25****
Chemotherapy	(0.026)	***(0.083)***	***(0.134)***	***(0.188)***	***(0.206)***	***(0.097)***
Chemotherapy vs	0.02	***0.17****	0.01	-0.10	0.04	0.02
Chemotherapy & Radiation	(0.025)	***(0.081)***	(0.132)	(0.184)	(0.201)	(0.095)
Radiation vs	-0.01	-0.03	***-0.23*****	***-0.42*****	***-0.48*****	***-0.23*****
Chemotherapy & Radiation	(0.017)	(0.054)	***(0.089)***	***(0.124)***	***(0.136)***	***(0.064)***

Note: Bold font indicates significance; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 3.13: Number (percent) of study patients in each treatment category.

Treatment	$N = 189$
Chemotherapy only	20 (10.6%)
Radiation only	64 (33.9%)
Both chemotherapy and radiation	77 (40.7%)
Neither chemotherapy nor radiation	28 (14.8%)

CHAPTER 4

Investigating the factor-analysis structure of the PAOFI data

Motivated by the Mind-Body Study (MBS), we are especially interested to know how differences on dimensions of the PAOFI might emerge across alternative primary post-surgical breast-cancer treatments (chemotherapy only, radiation only, both chemotherapy and radiation, neither chemotherapy nor radiation). In this chapter, we propose exploratory factor analysis to understand the effect of alternative scoring procedures on the number and composition of underlying factors, and we explore the between-treatment-group differences within the emerging factors. We now elaborate on each of these ideas in turn.

4.1 Motivation

The motivation of this investigation is based in part on experience in other contexts where the composite-variable structure that is widely used for a multi-item instrument is based on a factor analysis in a particular population, and there might be good reason to question whether the same composite variables would emerge from a factor analysis in a different population.

4.2 Literature review

The Patient's Assessment of Own Functioning Inventory (PAOFI) is a self-report QOL instrument designed by Chelune and colleagues as a tool for studying the relationship between patients' objective neuropsychological (NP) measures and their self-perceived capabilities and limitations in everyday living activities (Chelune et al., 1986). While the original instruments consisted of 47 items, Chelune et. al. utilized a factor analysis technique on a set of 33 ordinally-scaled questionnaire items, which resulted in extraction of 4 domains of cognitive functioning. A modified version of the PAOFI is a questionnaire consisting of 33 items that span four domains of cognitive functioning: memory (items $1-10$; e.g., "How often do you forget events which have occurred in the last day or two?"), language and communication (items $11-19$; e.g., "How often do you have difficulty thinking of names of things?"), sensorimotor abilities (items $20-24$; e.g., "How often do you have difficulty feeling things with your right hand?") and higher-level cognition (items $25-33$; e.g., "How often do you have difficulty finding your way about?"). Each item elicits a response on a 6-point Likert-type scale regarding how often a patient has recently been experiencing a particular type of disability in their everyday life, with 1 indicating "almost always" and 6 corresponding to "almost never".

The PAOFI has been used to elicit self-reports of cognitive impairments in a variety of health-research settings. In breast cancer research, Van Dyk and colleagues explored the factor structure of the PAOFI using the Mind Body Study sample of breast-cancer survivors where "chemobrain" phenomenon presents challenges for quality of life. The instrument was reported to measure 5 cognitive domains: higher-level cognition (HLC), which accounted for most of the variance in the data, memory-absent-mindedness, memory-forgetfulness, language production, and motor/sensory-perceptual (Van Dyk et al., 2016). As can be seen from Table 4.1, there is substantial overlap but not perfect agreement between the number and composition of the emerging factors in analyses by Chelune and colleagues versus those

from Van Dyk et al.

In the next section we reanalize data from the MBS from the perspective of looking at different number of factors and different cut-points.

4.3 Five-factor solutions

Focusing first on five-factor solutions to the factor-analysis problem, Table 4.2 illustrates the extent to which different ways of incorporating PAOFI data (i.e., using the original 6-point Likert scale, Q, or using one of the possible alternatives for dichotomizing the ordinal responses, $W^{(2)}$-$W^{(4)}$) yields similar or different groupings of items. The rows of Table 4.2 correspond to the 33 items in the PAOFI, and the columns correspond to different approaches that could be used to develop composite variables based on PAOFI items. The determination of which items should be grouped together in the columns is based on obtaining a five-factor solution using principal-component factor analysis, performing varimax rotation (to favor factor loadings near 0 or 1 in absolute value), and identifying for each item the factor on which the absolute value of the factor loading is the largest.

The entries in Table 4.2 refer to the factor number on which the given item loaded most highly. One could think of grouping the items that loaded most highly on a given factor by summing them or averaging them. Such a strategy could be interpreted as approximating the loadings for the items that loaded most highly on the given factor as 1 or -1 and approximating the loadings for the items that loaded more highly on another factor as 0. Composite variables could then be characterized by summing or averaging the items that load most highly on a given factor. The statistical methods being compared are: (1) the Chelune et al. approach that used data on the original ordinal scale (labeled "Chelune (1986)"), (2) the Van Dyk et al. (2016) approach that used data on the original ordinal scale and excluded items where the highest factor loading on any factor was less than 0.5 (labeled "Van Dyk (2016)"), (3) entering and analyzing the data on the original ordinal scale (labeled "Q"), (4)

an approach where the data entered into the factor analysis were dichotomized versions of PAOFI data with the cut point between 1/2 and 3/4/5/6 (labeled $W^{(2)}$), (5) factor analysis of dichotomized PAOFI data where the cut point is between 1/2/3 and 3/4/5/6 (labeled $W^{(3)}$), and (6) factor analysis of dichotomized PAOFI data where the cut point is between 1/2/3/4 and 5/6 (labeled $W^{(4)}$). Although there is a subjective and arguably ad hoc element to translating items into composite variables in this manner, where many factor loadings are implicitly set to zero so that the resulting linear combinations of items will reflect greater simplicity in the constructs being measured, the resulting simplification can help facilitate interpretation of complicated multivariate data.

A key question raised by this analysis is the extent to which different statistical methods give rise to similar or different combinations of items, corresponding to the question of whether analyses of quality-of-life data might be sensitive to choices in how factor analysis is implemented. We can also anticipate an accompanying benefit of listing results by factor number, as the factors are ordered in terms of the proportion of variance in the multivariate data explained by the given factor. Thus, arraying the findings from different statistical methods side-by-side, we will be able to observe not only the extent to which items are grouped together similarly across the different analysis methods but also the extent to which the factors on which those items are loading are explaining similar amounts of variability in the data.

As can be seen by the entries in Table 4.2, there is a nearly perfect agreement between the factor number and the corresponding items from Van Dyk et al. and Q, with the exception of the 3 items excluded by Van Dyk and colleagues due to item loadings being less than 0.5. Furthermore, a number of identical item groupings persist across multiple data treatments. For example, items 1, 2, 4, and 5 cluster into a factor in Q, $W^{(2)}$, $W^{(3)}$, and $W^{(4)}$; items 20 through 24 form the same factor in Q, $W^{(3)}$, and $W^{(4)}$. However, there are more differences than similarities in the resulting factor structure across the various data-scoring schemes. In addition to the emerging factors being comprised of non-identical combinations of items,

similar item groupings are not consistent in the order of the amount of variance they explain from Q to $W^{(4)}$. For example, items 1, 2, 4, 5 appear to form factor 3 in Q, meaning that these items account for third largest amount of variance explained unaccounted for by the first 2 combinations of items. In $W^{(2)}$, this combination of items loads on factor 2; whereas in $W^{(3)}$, the same items load on factor 4.

Of the 10 PAOFI items related to memory (Memory1-Memory10), both Chelune et al. and Van Dyk et al. found that the first 4 items (items 1-4) loaded on one factor and another 4 (items 6, 7, 9, 10) of the remaining items loaded on another factor. This structure was mirrored using approach $W^{(2)}$, but not $W^{(3)}$ and $W^{(4)}$.

While memory items 1-4 stay together in almost all approaches, items 10-19, identified by Chelune as 'language' items, are distributed on a variety of factors for all other methods.

4.4 Four-factor solutions

Building on the previous analysis, Table 4.3 displays for each of several factor-analysis strategies the factor number on which each item loads most highly, where in contrast to Table 4.2 which provided findings based on five-factor solutions, the entries in Table 4.3 are based on four-factor solutions. As before, the statistical methods being compared are entering and analyzing the data on the original ordinal scale (labeled Q), and three analyses where the data entered into the factor analysis were dichotomized versions of PAOFI data, where the cut point was either between 1/2 and 3/4/5/6 (labeled $W^{(2)}$), between 1/2/3 and 3/4/5/6 (labeled $W^{(3)}$), or between 1/2/3/4 and 5/6 (labeled $W^{(4)}$). The entries in Table 4.3 are analogous to those in Table 4.2, but for a 4-factor solution instead of 5-factor solution.

As can be seen by the entries in Table 4.3, several item clusters consistently load on the same factors across all data-treatment variations, Q through $W^{(4)}$. For example, items 1-5, all of which relate to memory, load on a single factor in all 4 cases; items 30-33, which relate to HLC, also load on the same factor across Q through $W^{(4)}$; items 20-23 load on the same

factor across all cases except in $W^{(2)}$.

However, there are also notable differences across strategies. For example when we look at items 11-19 which have a language related context, we see considerable variation across other methods. For example, in method $W^{(2)}$. Thus, similar to a 5-factor solution, we see that there is no consistency in the order of the emerging factors, nor is there a unique linear combination preserved for each factor from Q to $W^{(4)}$.

4.5 Six-factor solutions

Table 4.4 displays factor number on which each item loads most highly for each of the given data coding strategy (Q, $W^{(2)}$ to $W^{(4)}$). As before, the statistical methods being compared are entering and analyzing the data on the original ordinal scale (labeled Q), and three analyses where the data entered into the factor analysis were dichotomized versions of PAOFI data, where the cut point was either between 1/2 and 3/4/5/6 (labeled $W^{(2)}$), between 1/2/3 and 3/4/5/6 (labeled $W^{(3)}$), or between 1/2/3/4 and 5/6 (labeled $W^{(4)}$). The entries in Table 4.4 are analogous to those in Table 4.2, but for a 6-factor solution instead of 5-factor solution.

Of the 10 PAOFI items related to memory (Memory1-Memory10), items 1-5 loaded on the same factor in all approaches except $W^{(3)}$. Items 6, 7, 9, 10 loaded on the same factor in Q and $W^{(2)}$, but distributed on a variety of factors in $W^{(3)}$ and $W^{(4)}$.

Of the 9 items related to language (Language1-Language9), items 16-19 loaded on the same factor in Q, items 14-19 (but not item 18) loaded on the same factor in $W^{(3)}$. Other items dispersed over multiple factors in all data-coding strategies.

Of the 5 items related to sensorimotor function, items 20-23 loaded on the same factor in Q, $W^{(3)}$ and $W^{(4)}$, but not in $W^{(2)}$. In $W^{(3)}$ it was factor 3, while in both Q and $W^{(3)}$ it was factor 4.

Of the 9 items related to higher-level cogntion (HLC), all items loaded on the same factor, factor 1, in Q and $W^{(4)}$. Items 30-33 loaded on the same factor in all data-scoring strategies.

4.6 Comparisons of breast-cancer treatments based on factor scores

Of particular interest in the present context is to contrast outcomes on PAOFI factors across different breast-cancer treatments. In the Mind-Body Study, participants were grouped based on whether they were treated with both chemotherapy and radiation, chemotherapy alone, radiation alone, or neither. Here, we investigate how alternative factor-analysis strategies might affect the understanding of the extent to which patients have different quality of life across different domains of experience. Reversing the coding of outcomes as needed so that higher item scores imply worse quality of life, we consider three factor-analysis strategies: the original approach taken by Chelune et al. (1986) (using a four-factor solution where memory items were pooled), a five-factor solution with ordinal item scoring (distinguishing forgetfulness and absent-mindedness in the memory domain in line with Van Dyk et al. (2016)), and a five-factor solution with binary item scoring (again distinguishing forgetfulness and absent-mindedness). The items contributing to the respective factors have substantial overlap but are not identical; nonetheless, we view the comparison as being of interest, as different investigators might associate the same labels to factors (e.g., "Memory: Forgetfulness" or "Language") based on substantially overlapping content. The factor analysis reported by Van Dyk et al. (2016) was similar to what is reported here as the five-factor solution with ordinal item scoring, with the difference being that Van Dyk et al. omitted three items that had factor loadings below 0.5 on all factors, while the approach here includes those items on the factor where they had the highest loading.

For each alternative factor-analysis strategy, we calculated composite outcome measures. Table 4.5 reports findings based on performing a one-way analysis of variance and carrying out all pairwise comparisons between group means using the Tukey procedure to account

for multiple comparisons in a way that bounds the experiment-wise error rate for the pairwise comparisons at 0.05. Table 4.6 reports the corresponding findings based on Wilcoxon rank-sum comparisons among all pairs of breast-cancer treatment arms and incorporating a Bonferroni adjustment for there being 6 pairwise comparisons among the group means.

As might have been anticipated, there were distinctions in the findings across the alternative factor-analysis strategies. For example, with ordinal item scoring, significant differences not detected with the other factor-analysis strategies were seen on the Language factor between chemotherapy alone and radiation alone and between the group with both chemotherapy and radiation and the group with neither chemotherapy nor radiation. However, as also might have been anticipated, there were certain findings that were robust to the choice of factor-analysis strategy. For example, all three strategies found a significant difference on the Language factor with higher scores on average for the combined chemotherapy-and-radiation group compared with radiation alone. The significance of the latter difference did not depend on the statistical analysis strategy, as the same finding was seen both in Table 4.5 using one-way ANOVA with a Tukey multiple-comparison procedure and in Table 4.6 with Wilcoxon rank-sum tests supplemented with a Bonferroni adjustment for multiple comparisons.

4.7 Item-Response Theory modeling of PAOFI Data

Another multivariate modeling approach for analyzing PAOFI data is to use item-response-theory (IRT) models (Hays et al., 2000). The underlying idea is to conceptualize scores from a multi-item instrument as arising from a process where, on a logistic scale, additive effects across the levels of an item are combined with an additive person effect to account for the frequency with which the levels of a given item are endorsed. Using the software available in SAS PROC IRT, it is possible to gain insight into how individual items contribute to factors comprised of those items. Specifically, we fit generalized partial credit model (GPCM) which characterizes the probability, P_{ij}, of endorsing the j^{th} category from m possible categories if

item i given the latent characteristic, θ, in the following way:

$$P_{ij}(\theta) = \frac{exp(\sum_{k=1}^{r}[a_i(\theta - b_{ik})])}{1 + \sum_{r=1}^{m}(exp\sum_{k=1}^{r}[a_i(\theta - b_{ik})])} \tag{4.1}$$

In this formula, a_i is a slope parameter that helps characterize the ability of an item to discriminate between people of various ability on the underlying latent construct, θ. Probabilities associated with endorsing different levels of the item given θ are visualized using item characteristic curves, where b_{ik} indicates the level of θ where adjacent probability curves intersect. An appealing feature of the item characteristic curves is to have well-separated probability distributions over the underlying latent variable, which is indicative of the item being informative about individuals.

Building on the previously reviewed factor analyses, the strategy used here was to fit IRT models separately for each factor. Specifically, we focused attention on the five-factor solution where PAOFI data were entered on the original ordinal scale and where, unlike Van Dyk et al. (2016), we retained all PAOFI items in the factor analysis rather than excluding any items. As noted in Section 4.3, the resulting factors were similar to those identified in the Van Dyk et al. investigation.

In Figure 4.1 through Figure 4.5, we present item-characteristic curves emerging from the respective fitted IRT models. Differently colored curves reflect relative likelihood of endorsing particular level of item given θ. Quantitative reflections of the information in these figures are provided by the estimated slopes across the levels of each item. For Figure 4.1 (items 1, 2, 3, 4, 5, and 14), the estimated slopes are 2.25, 2.83, 2.92, 2.05, 1.57, and .99. For Figure 4.2 (items 6, 7, 9, 10, and 26), the estimated slopes are 2.01, 2.17, 1.62, 2.16, and 1.35. For Figure 4.3 (items 15, 16, 17, 18, and 19), the estimated slopes are .84, 2.94, 4.10, 1.28, and 1.20. For Figure 4.4 (items 20, 21, 22, and 23), the estimated slopes are 1.05, 1.36, 3.61, and 2574. For Figure 4.5 (items 8, 11, 12, 13, 24, 25, 27, 28, 29, 30, 31, 32, and 33), the estimated slopes are 1.62, 1.84, 1.86, 2.32, 1.26, 2.41, 2.23, 1.78, 2.35, 2.70, 3.59, 3.27,

3.82.

Qualitatively, items that contribute meaningfully to a given factor appear as having components that are well-separated from other components. For example, in the factor comprised of items 15, 16, 17, 18, and 19, the representation of items 16 and 17 stand out as having well-separated components. We return to this perspective in Section 4.8 when considering the possibility of a short-form version of the PAOFI.

4.8 Exploration of possible short form of PAOFI

Given that the PAOFI has 33 items, and given the available evidence from factor analysis and IRT analysis of study findings, a question emerged as to whether it would be possible to recognize differences in quality-of-life outcomes across breast-cancer treatments with a smaller number of items. To this end, we considered a possible short-form version of the PAOFI comprised of 10 items by including two items from each of five factors.

In fitting the GPCM we obtained estimates of the slopes as well as graphical displays of the implied ability to discriminate between levels of items. After producing these item characteristic curves for each factor we selected 2 items for each factor based on subjective assessment of item characteristic curves and the slopes of the items. This led to our choosing the following 2 items for each scale: first Memory factor: Item 3 ("Forgetting people whom you met in the last day or two") and Item 4 ("Forgetting things that you knew a year or more ago"); second Memory factor: Item 9 ("Losing things or having trouble remembering where they are") and Item 10 ("Forgetting things that you are supposed to do or have agreed to do"); Language factor: Item 16 ("Having difficulty thinking of the names of things") and Item 17 ("Having difficulty thinking of the words (other than names) for what you want to say"); Sensorimotor factor: Item 21 ("Having difficulty performing tasks with your left hand") and Item 22 ("Having difficulty feeling things with your right hand"); Higher-Level Cognition factor: Item 25 ("Having thoughts that seem confused or illogical") and Item 31

("Having more difficulty now than you used to in solving problems that come up around the house, at your job, etc."). We then re-analyzed the data using ANOVA supplemented with Tukey-adjusted pairwise comparisons (Table 4.7) and with nonparametric analogs (Table 4.8).

Table 4.7 presents the results from 1-way ANOVA from both the ordinal item 5-factor scoring and the composite variables based on 2 items per factors. Table 4.8 presents the similar side-by-side comparison for using Wilcoxon rank-sum test with Bonferroni adjustment for multiple comparisons to do pairwise comparisons between breast cancer treatments. In Table 4.7 we see that the reduced instrument that uses 2 items per factor is still able to detect most of the same differences across breast cancer treatment groups. In fact, it detects an additional one on the Language domain but does not pick up all of the same significant differences. For example, it did not detect significantly higher Memory: Absent-Mindedness score for study participants in the chemo and radiation group versus those with neither treatment.

Results from non-parametric analysis are analogous to those from 1-way ANOVA. In Table 4.7 we see that the reduced instrument that uses 2 items per factor is able to detect all but one significant comparisons for Memory: Absent-Mindedness, Language, and Higher-Level Cognition identified with using the full composite. Furthermore, it identifies 2 additional significant comparisons for Language, and an extra one for Higher-Level Cognition.

From these analyses, we maintained the ability to detect significant differences across breast-cancer treatments for the following factors: Memory: Absent-Mindedness, Language, and Higher-Level Cognition. No significant comparisons were identified for Memory: Forgetfulness and Sensorimotor.

4.9 Discussion

Important scientific considerations in any context are the degree to which research findings are replicable and the extent to which findings are robust to perturbations of underlying assumptions. In this chapter, we saw that alternative approaches to factor analysis of PAOFI measures gave rise to subtle distinctions in associated groupings of items. Such distinctions are not particularly surprising, as there are subjective ingredients in factor analysis procedures associated with ambiguities such as the equivalent fit to the data provided by any rotation of a factor-analysis solution.

Yet we also saw meaningful robustness of certain findings when comparing composite-variable outcomes across breast-cancer treatment groups. We explored summaries from a four-factor solution using ordinal-item scoring, a five-factor solution using ordinal-item scoring, and a five-factor solution using binary-item scoring. Across all of these approaches, certain contrasts in quality-of-life outcomes were seen to be statistically significant, offering robust insight that especially highlights the quality-of-life challenges facing women who undergo both chemotherapy and radiation treatment for breast cancer.

In one instance, there was a significant difference using ANOVA / Tukey-adjusted comparison for the factor-analysis approaches using ordinal-item scoring that was not seen using binary-item scoring, namely that the combined chemotherapy-and-radiation group had higher scores (i.e., worse outcomes) on the Higher-Level Cognition factor than the radiation-alone group. Reflecting on material presented earlier in the dissertation regarding the loss of information associated with dichotomizing ordinal or continuous measures, the finding that there was less ability to detect a significant difference across breast-cancer treatment groups using binary outcome scoring in factor analysis might be regarded as a further indication that dichotomizing measures is not to be encouraged when it is not essential.

A further question of potential interest is whether it would be possible to detect meaningful quality-of-life differences across breast-cancer treatment arms using a shorter instrument

that would entail less response burden for research participants. We considered this possibility in the context of an item-response-theoretic framework, using item characteristic curves and discrimination parameter estimates from fitted IRT models to create reduced composite variables based on PAOFI factor constructs each composed of 2 items.

ANOVA / Tukey-adjusted comparison for the factor-composition approaches revealed that the reduced instrument that uses 2 items per factor was still able to detect most, but not all, of the same differences across breast cancer treatment groups. With 1-way ANOVA significance findings were the same for Memory: Absent-Mindedness, Language, and Higher-Level Cognition domains. The reduced composite variable detected an additional significant comparison for the Language domain. Results from Wilcoxon rank-sum test were similar to those from 1-way ANOVA. While it should be expected that the results would not be identical with full factor composite variables, these findings support the possibility of employing a simplified version of the instrument whenever response burden for research participants, or time, or budget constraints are of concern.

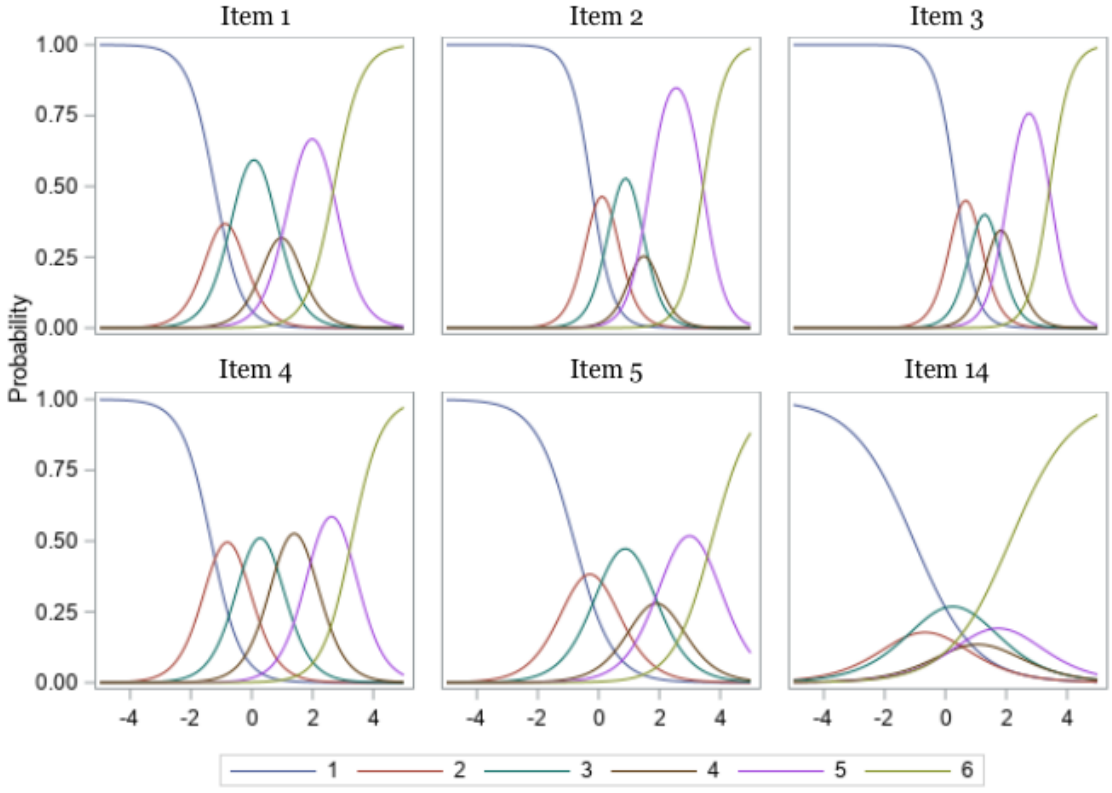

Figure 4.1: Item characteristic curves of items from the Memory: Forgetfulness domain.

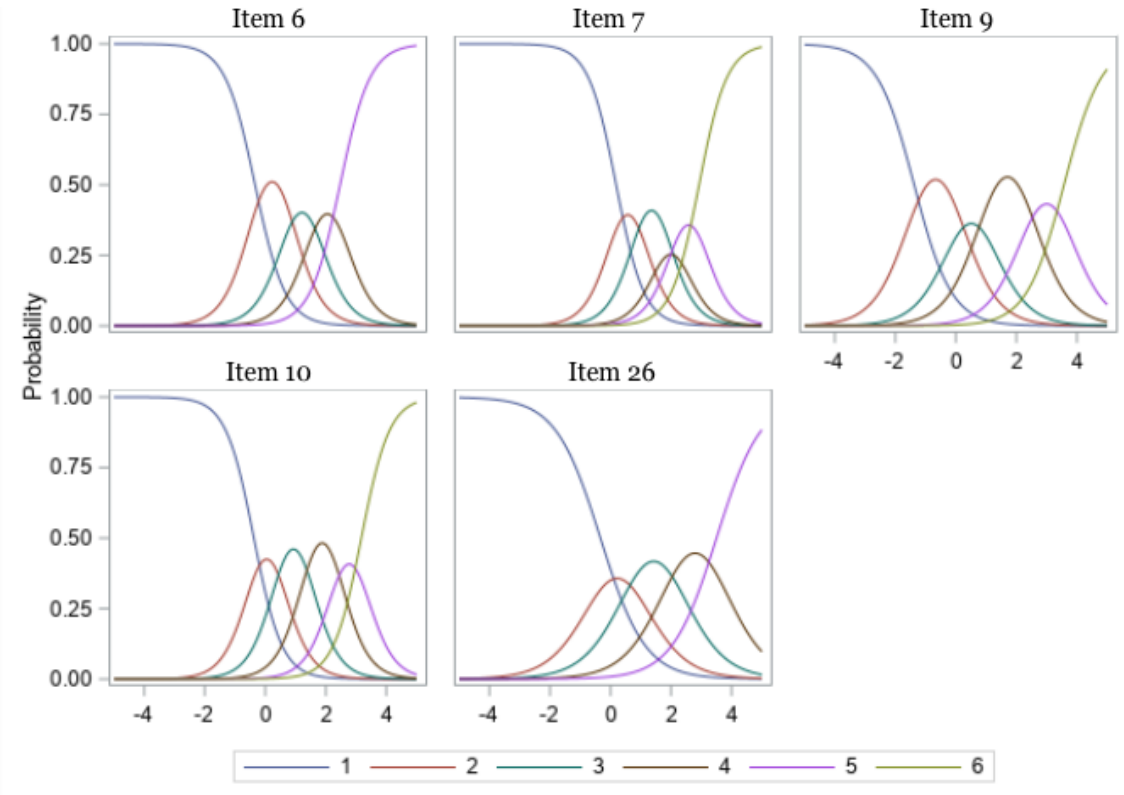

Figure 4.2: Item characteristic curves of items from the Memory: Absent-Mindedness domain.

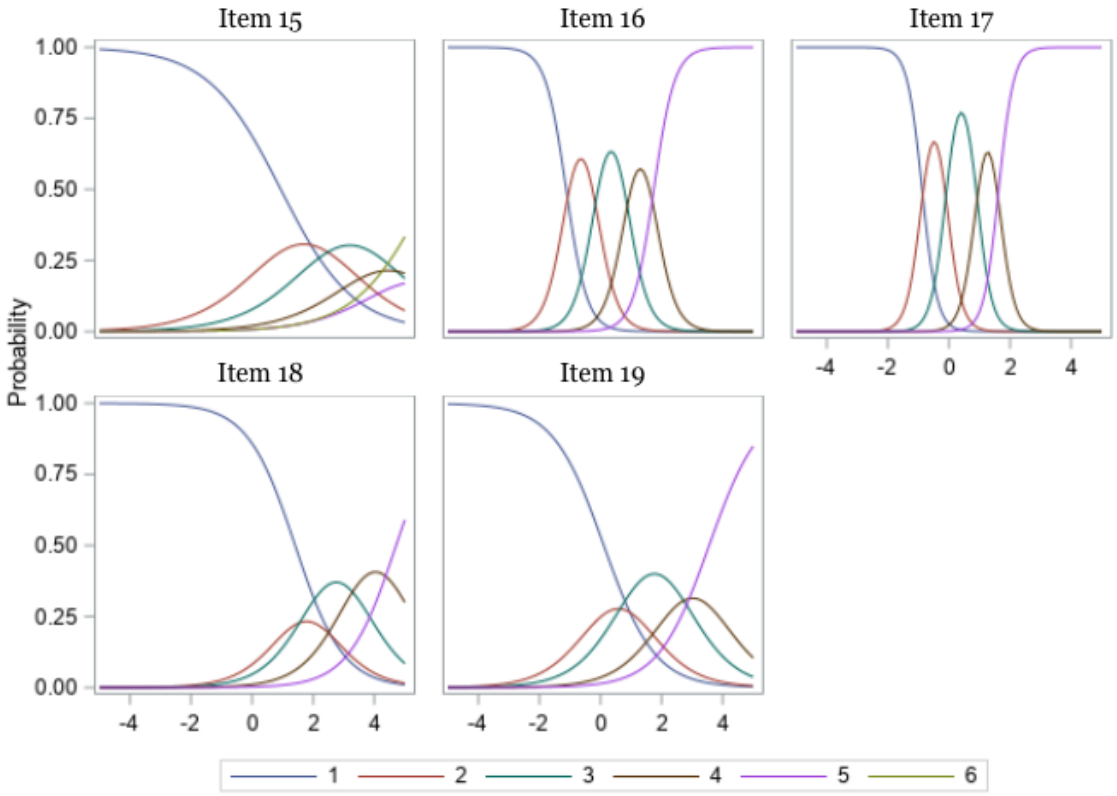

Figure 4.3: Item characteristic curves of items from the Language domain.

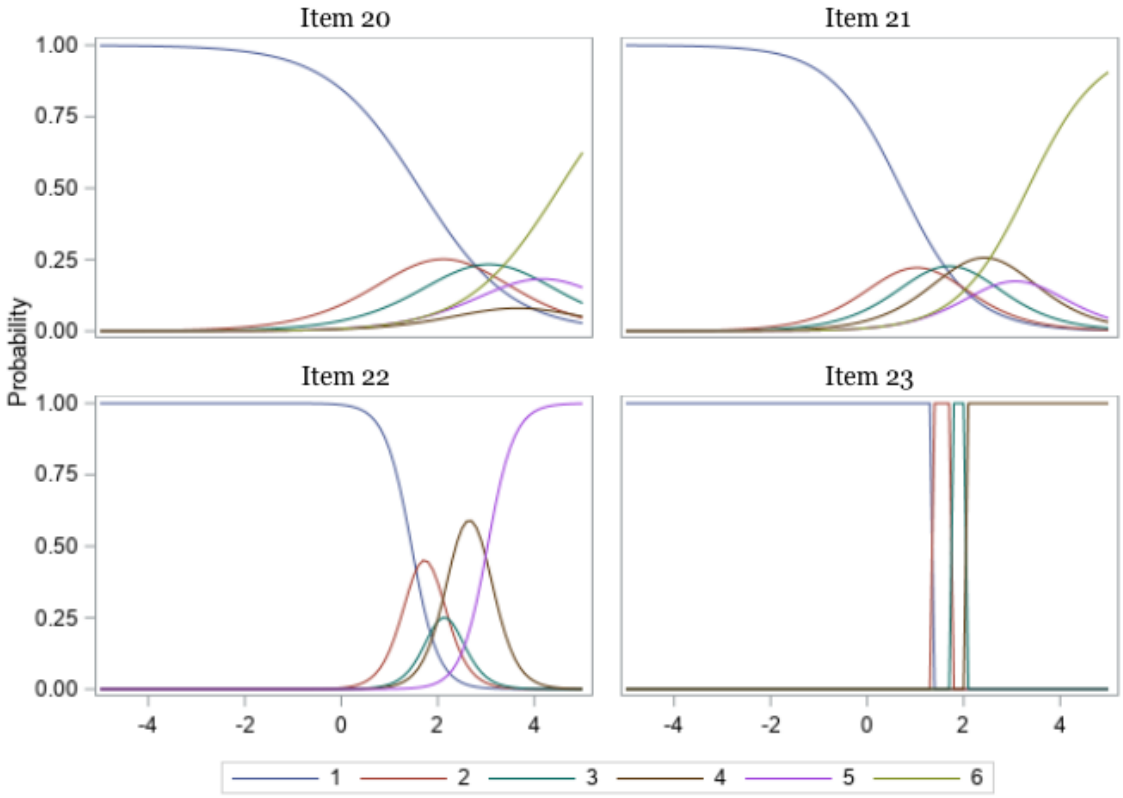

Figure 4.4: Item characteristic curves of items from the Sensorimotor domain.

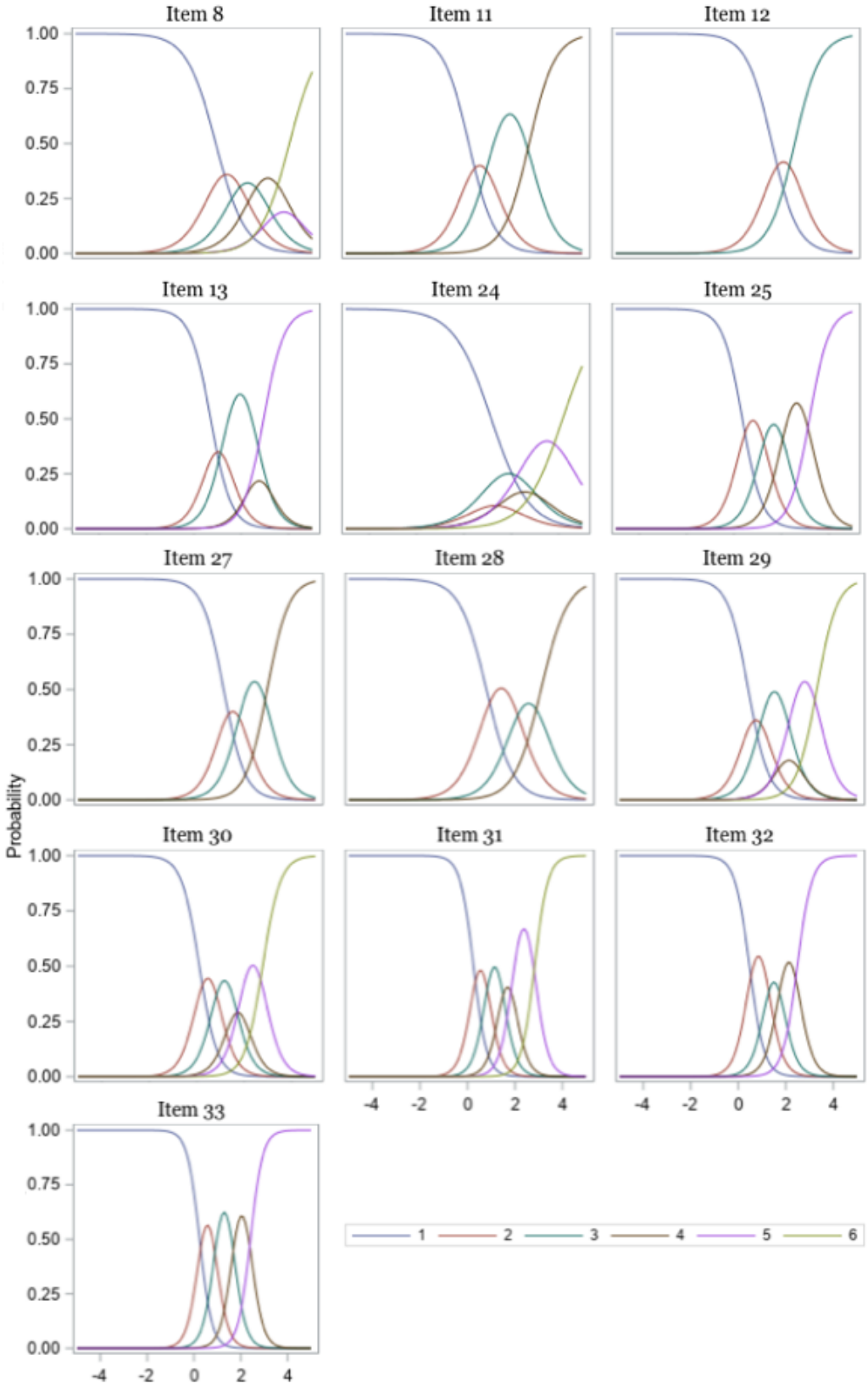

Figure 4.5: Item characteristic curves of items from the Higher-Level Cognition domain.

Table 4.1: PAOFI item descriptions and rotated factor structure from the studies by Chelune et al. (1986) and Van Dyk et al. (2016)

Item	Item description	Subject-matter grouping	Chelune (1986)	Van Dyk (2016)
1	Forgetting something that has been told to you within the last day or two	Mem1	Mem-Forget1	Mem-Forget1
2	Forgetting events which have occurred int he last day or two	Mem2	Mem-Forget2	Mem-Forget2
3	Forgetting people whom you met in the last day or two	Mem3	Mem-Forget3	Mem-Forget3
4	Forgetting things that you knew a year or more ago	Mem4	Mem-Forget4	Mem-Forget4
5	Forgetting people whom you knew or met a year or more ago	Mem5	Excl[1]	Mem-Forget5
6	Losing track of time, or do things either earlier or later than they are usually done	Mem6	Mem-Absent6	Mem-Absent1
7	Failing to finish something you start because you forgot that you were doing it	Mem7	Mem-Absent7	Mem-Absent2
8	Failing to complete a task because you have forgotten how to do aspects of it	Mem8	Excl[1]	HLC1
9	Losing things or having trouble remembering where they are	Mem9	Mem-Absent9	Mem-Absent3
10	Forgetting things that you are supposed to do or have agreed to do	Mem10	Mem-Absent10	Mem-Absent4
11	Having difficulties understanding what is said to you	Lang1	Lang1	HLC2
12	Having difficulties recognizing or identifying printed words	Lang2	Lang2	HLC3
13	Having difficulty understanding reading material which you formerly could have understood	Lang3	Lang3	HLC4
14	Easier to have people show you things than it is to have them tell you about things	Lang4	Lang4	Excl[2]
15	Having indistinct or improperly pronounced words when you speak	Lang5	Lang5	Excl[2]
16	Having difficulty thinking of the names of things	Lang6	Lang6	Lang1
17	Having difficulty thinking of the words (other than names) for what you want to say	Lang7	Lang7	Lang2
18	Having difficulty forming the letters correctly when you write things	Lang8	Lang8	Lang3
19	Having more difficulty spelling, or make more errors in spelling, than you used to	Lang9	Lang9	Lang4
20	Having difficulty performing tasks with your right hand	Hands1	Excl[1]	SM1
21	Having difficulty performing tasks with your left hand	Hands2	SM1	SM2
22	Having difficulty feeling things with your right hand	Percept1	SM2	SM3
23	Having difficulty feeling things with your left hand	Percept2	SM3	SM4
24	Having more difficulty than you used to in seeing all of what you are looking at	Percep1	SM4	Excl[2]
25	Having thoughts that seem confused or illogical	HLC1	HLC1	HLC5
26	Become distracted from what you are doing or saying by insignificant things	HLC2	HLC2	Mem-Absent5
27	Becoming confused about (or make a mistake about) where you are	HLC3	HLC3	HLC6
28	Having difficulty in finding your way about	HLC4	HLC4	HLC7
29	Having more difficulty now than you used to in calculating or working with numbers	HLC5	HLC5	HLC8
30	Having more difficulty now than you used to in planning or organizing activities	HLC6	HLC6	HLC9
31	Having more difficulty now than you used to in solving problems that come up around the house, at your job, etc.	HLC7	HLC7	HLC10
32	Having more difficulty now than you used to in following directions to get somewhere	HLC8	HLC8	HLC11
33	Having more difficulty now than you used to in following instructions on how to do things	HLC9	HLC9	HLC12

Notes: "Excl[1]" signifies items excluded from analysis for unspecified reasons; "Excl[2]" signifies items excluded from factors due to their loadings being less than 0.5.

Table 4.2: Comparison of rotated five-factor solutions using different rules for item scoring

Item	Subject-matter grouping	Chelune (1986)	Van Dyk (2016)	Q	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$
1	Memory1	4	3	3	2	4	3
2	Memory2	4	3	3	2	4	3
3	Memory3	4	3	3	2	1	3
4	Memory4	4	3	3	2	4	3
5	Memory5	Excl[1]	3	3	2	4	3
6	Memory6	2	2	2	3	1	2
7	Memory7	2	2	2	3	4	3
8	Memory8	Excl[1]	1	1	3	1	1
9	Memory9	2	2	2	3	4	4
10	Memory10	2	2	2	3	4	3
11	Language1	1	1	1	Excl[3]	2	2
12	Language2	1	1	1	Excl[3]	Excl[3]	1
13	Language3	1	1	1	2	2	1
14	Language4	1	Excl[2]	3	2	5	3
15	Language5	1	Excl[2]	4	4	5	5
16	Language6	1	4	4	5	5	4
17	Language7	1	4	4	5	5	4
18	Language8	1	4	4	1	3	5
19	Language9	1	4	4	1	5	1
20	Sensorimotor1	5	5	5	5	3	5
21	Sensorimotor2	5	5	5	4	3	5
22	Sensorimotor3	5	5	5	5	3	5
23	Sensorimotor4	5	5	5	Excl[3]	3	5
24	Sensorimotor5	Excl[1]	Excl[2]	1	1	2	1
25	HLC1	3	1	1	4	1	2
26	HLC2	3	2	2	4	1	2
27	HLC3	3	1	1	Excl[3]	3	1
28	HLC4	3	1	1	Excl[3]	3	1
29	HLC5	3	1	1	2	1	1
30	HLC6	3	1	1	1	2	2
31	HLC7	3	1	1	5	2	2
32	HLC8	3	1	1	1	2	1
33	HLC9	3	1	1	1	2	2

Notes: "Excl[1]" signifies items excluded from analysis for unspecified reasons; "Excl[2]" signifies items excluded from factors due to their loadings being equal to less than 0.5; "Excl[3]" signifies items excluded from factors due to their loadings being equal to 0.

Table 4.3: Comparison of rotated four-factor solutions using different rules for item scoring

Item	Subject-matter grouping	Q	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$
1	Memory1	3	3	1	2
2	Memory2	3	3	1	2
3	Memory3	3	3	1	2
4	Memory4	3	3	1	2
5	Memory5	3	3	1	2
6	Memory6	2	2	1	2
7	Memory7	2	4	1	2
8	Memory8	1	2	2	1
9	Memory9	2	2	1	3
10	Memory10	2	2	1	2
11	Language1	1	Excl[1]	2	1
12	Language2	1	Excl[1]	Excl[1]	1
13	Language3	1	3	2	1
14	Language4	3	3	4	2
15	Language5	4	4	4	4
16	Language6	2	2	4	3
17	Language7	2	2	4	3
18	Language8	4	1	3	4
19	Language9	2	1	4	1
20	Sensorimotor1	4	2	3	4
21	Sensorimotor2	4	4	3	4
22	Sensorimotor3	4	4	3	4
23	Sensorimotor4	4	Excl[1]	3	4
24	Sensorimotor5	1	1	2	1
25	HLC1	1	4	1	1
26	HLC2	1	4	1	1
27	HLC3	1	Excl[1]	3	1
28	HLC4	1	Excl[1]	3	1
29	HLC5	1	3	2	1
30	HLC6	1	1	2	1
31	HLC7	1	1	2	1
32	HLC8	1	1	2	1
33	HLC9	1	1	2	1

Note: "Excl[1]" signifies items excluded from factors due to their loadings being equal to 0.

Table 4.4: Comparison of rotated six-factor solutions using different rules for item scoring

Item	Subject-matter grouping	Q	$W^{(2)}$	$W^{(3)}$	$W^{(4)}$
1	Memory1	3	2	2	2
2	Memory2	3	2	2	2
3	Memory3	3	2	1	2
4	Memory4	3	2	6	2
5	Memory5	3	2	6	2
6	Memory6	2	3	1	1
7	Memory7	2	3	2	4
8	Memory8	1	3	3	3
9	Memory9	2	3	2	4
10	Memory10	2	3	2	4
11	Language1	1	Excl[1]	3	1
12	Language2	1	Excl[1]	Excl[1]	5
13	Language3	1	2	3	1
14	Language4	3	2	5	2
15	Language5	6	5	5	5
16	Language6	4	4	5	4
17	Language7	4	4	5	4
18	Language8	4	1	4	5
19	Language9	4	1	5	5
20	Sensorimotor1	5	4	6	6
21	Sensorimotor2	5	6	6	5
22	Sensorimotor3	5	6	4	6
23	Sensorimotor4	5	Excl[1]	4	6
24	Sensorimotor5	1	1	3	5
25	HLC1	2	5	1	1
26	HLC2	2	5	1	1
27	HLC3	1	Excl[1]	4	3
28	HLC4	1	Excl[1]	4	3
29	HLC5	1	2	1	1
30	HLC6	1	5	1	1
31	HLC7	1	4	3	1
32	HLC8	1	1	3	1
33	HLC9	1	1	3	1

Note: "Excl[1]" signifies items excluded from factors due to their loadings being equal to 0.

Table 4.5: Significant contrasts in 1-way ANOVA across breast cancer treatment groups

	Factor analysis strategy		
Factor domain	Chelune subject-matter grouping (4 factors)	Ordinal item scoring (5 factors)	Binary item scoring (5 factors)
Memory	C&R > R alone C&R > Neither	N/A	N/A
Memory: Forgetfulness	N/A	None	C&R > R alone
Memory: Absent-mindedness	N/A	C&R > R alone C&R > Neither C alone > R alone	C&R > R alone C&R > Neither C alone > Neither
Language	C&R > R alone	C alone > R alone C&R > R alone C&R > Neither	C&R > R alone
Sensorimotor	None	None	None
Higher-level cognition	C&R > Neither C&R > R alone	C&R > Neither C&R > R alone	C&R > Neither

Notes:

Factor domains are based on rotated factor solutions.

Table entries refer to contrasts significant at .05 level.

"A>B" implies mean of group A greater than mean of group B.

"C alone": chemotherapy alone

"R alone": radiation alone

"C&R": chemotherapy and radiation

"Neither": neither chemotherapy nor radiation

Table 4.6: Significant contrasts in Wilcoxon rank-sum test across breast cancer treatment groups

	Factor analysis strategy		
Factor domain	Chelune subject-matter grouping (4 factors)	Ordinal item scoring (5 items)	Binary item scoring (5 factors)
Memory	C&R > R alone C&R > Neither	N/A	N/A
Memory: Forgetfulness	N/A	None	C&R > R alone
Memory: Absent-mindedness	N/A	C&R > Neither C&R > R alone	C&R > Neither
Language	C&R > R alone	C&R > R alone C&R > Neither	C&R > R alone
Sensorimotor	None	None	None
Higher-level cognition	C&R > R alone C&R > Neither	C&R > R alone C&R > Neither	C&R > R alone C&R > Neither

Notes:

Factor domains are based on rotated factor solutions.

Table entries refer to contrasts significant at .05 level.

"A>B" implies mean of group A greater than mean of group B.

"C alone": chemotherapy alone

"R alone": radiation alone

"C&R": chemotherapy and radiation

"Neither": neither chemotherapy nor radiation

Table 4.7: Significant contrasts in 1-way ANOVA across breast cancer treatment groups

Factor domain	Analysis strategy	
	Full composite (based on all items)	Reduced composite (based on 2 items/factor)
Memory: Forgetfulness	None	None
Memory: Absent-mindedness	C&R > R alone C&R > Neither C alone > R alone	C&R > R alone
Language	C&R > R alone C&R > Neither C alone > R alone	C&R > R alone C&R > Neither C alone > R alone C alone > Neither
Sensorimotor	None	None
Higher-level cognition	C&R > R alone C&R > Neither	C&R > R alone C&R > Neither

Notes:

Factor domains are based on rotated factor solutions.

Table entries refer to contrasts significant at .05 level.

"A>B" implies mean of group A greater than mean of group B.

"C alone": chemotherapy alone

"R alone": radiation alone

"C&R": chemotherapy and radiation

"Neither": neither chemotherapy nor radiation

Table 4.8: Significant contrasts in Wilcoxon rank-sum test across breast cancer treatment groups

Factor domain	Analysis strategy	
	Full composite (based on all items)	Reduced composite (based on 2 items/factor)
Memory: Forgetfulness	None	None
Memory: Absent-mindedness	C&R > R alone C&R > Neither	C&R > R alone
Language	C&R > R alone C&R > Neither	C&R > R alone C&R > Neither C alone > R alone C alone > Neither
Sensorimotor	None	None
Higher-level cognition	C&R > R alone C&R > Neither	C&R > R alone C&R > Neither C alone > Neither

Notes:

Factor domains are based on rotated factor solutions.

Table entries refer to contrasts significant at .05 level.

"A>B" implies mean of group A greater than mean of group B.

"C alone": chemotherapy alone

"R alone": radiation alone

"C&R": chemotherapy and radiation

"Neither": neither chemotherapy nor radiation

CHAPTER 5

Three-Part Model for PAOFI Data

Motivated by the Mind-Body Study (MBS), we are especially interested to know how differences on dimensions of the PAOFI might emerge across alternative primary post-surgical breast-cancer treatments (chemotherapy alone, radiation alone, both chemotherapy and radiation, neither chemotherapy nor radiation). In this chapter, we propose a multi-part model that improves upon the widely-used dichotomization of the PAOFI by complementing such a dichotomization with additional analyses of subparts of the data.

5.1 Motivation

As noted in the introduction to this chapter, a three-part model improves upon the widely used dichotomization by complementing it with additional analyses of subparts of the data. In this section, we introduce notation for how such a development might proceed.

The following approach is motivated by the work of Duan et al. (1984) who applied two-part model to health services utilization data, and the framework developed by Bradlow and Zaslavsky (1999) in a context where ordinal data were obtained to evaluate customer satisfaction with the DuPont Corporation. Based on subject-matter considerations, it made sense in that context to develop a multi-part model where one stage involved using logistic regression to identify the factors of non-response, and among cases where the ratings were present, a next stage of the model fitted ordinal logistic regression to identify the predictors of customer satisfaction. Their approach was embedded in a Bayesian framework

with assumptions characterizing the prior distributions of parameters at different stages of a multi-part model.

Here, we start with the less ambitious goal of developing a reasonable three-part model for the data in the PAOFI setting. A three-part model has the advantage of being able to offer more insights into data, than a simple dichotomization approach that is widely used. In particular, it can discern the differences between patients who do and do not report cognitive dysfunction. It can also differentiate between patients reporting more complaints from those with more severe complaints. For example, suppose patient A has scores of $1, 4, 4$, and patient B has scores of $3, 3, 3$ on a 3-item subset of the PAOFI. As lower score corresponds to having difficulty more often, patient A seems to struggle considerably while performing tasks on 1 item, while patient B has some difficulty performing tasks on all 3 items. The simple dichotomization would mask the differences in the number and severity of complaints, while a three-part model would consider each of these aspects of measuring complaints separately.

We develop notation for the three-part model in the following way:

Q_{dij} = Response of person $i = 1, ..., n$ to PAOFI item $j = 1, .., J$ in domain $d = 1, ..., D$;

$Q_{di+} = \sum_{j=1}^{J} Q_{dij}$ = Sum of items in domain d for person i;

$\bar{Q}_d = \frac{1}{n} \sum_{i=1}^{n} Q_{di+}$ = Average of summed items in domain d;

$$W_{dij}^{(k)} = \begin{cases} 1 & \text{if } Q_{dij} \leq k \\ 0 & \text{if } Q_{dij} > k \end{cases} = \text{Dichotomized response at threshold } k = 1, ..., 5;$$

$W_{di+}^{(k)} = \sum_{j=1}^{J} W_{dij}^{(k)}$ = Sum of dichotomized items;

$\bar{W}_d^{(k)} = \frac{1}{n} \sum_{i=1}^{n} W_{di+}^{(k)}$ = Average of summed dichotomized items.

A potential three-part model may be modeled as follows:

1. Logistic model identifying $W_{di+}^{(k)} > 0$ vs $W_{di+}^{(k)} = 0$. This model may be used to identify predictors of having any issues vs none.

2. Truncated negative binomial model fit to $W_{di+}^{(k)} > 0$ subset. In those reporting issues, what factors predict having more vs fewer issues.

3. Linear regression model fit to $W_{di+}^{(k)} > 0$ subset. This part of the model may offer insights into the predictors of average severity of symptoms ($\bar{Q}_d$), or the aggregate impact of symptoms (Q_{di+}).

The plan for the dissertation is to illustrate the strategy using data from the Mind-Body Study using $\{1, 2, 3\}$ versus $\{4, 5, 6\}$ as a primary dichotomization. We first explore single-predictor models to gain insight into the roles of individual predictor variables. We then investigate multiple-predictor models where indicators for breast cancer treatment are forced in the model, thereby focusing attention on the extent to which breast-cancer-treatment comparisons are affected by the inclusion of other predictors as well as the extent to which other predictor variables explain variation in PAOFI outcomes after accounting for differences across breast cancer treatments.

5.2 Three-part models with individual predictor variables

Taking summaries from each PAOFI factor as outcome variables, we fit a sequence of models to implement the three-part-model strategy. With two variations for the third part of the model, with either average outcome values associated with the items where problems were noted or the sum of outcome values associated with the items where problems were noted, we report results from four models for each PAOFI factor domain.

Findings for the Memory: Forgetfulness domain are shown in Table 5.1. The first column of the table reports significant predictors from logistic regression models where the outcome is an indicator for any reported problems in the given domain. A variable is listed if its association with the outcome reached at least the 0.25 level of significance. The display is in italic font if the predictor reached the 0.10 level of significance, and the display is in italic

bold font if the predictor reached the 0.05 level of significance. For ordinal, continuous, and binary predictors, the variables listed in the table are appended with "(+)" if the association with the outcome is positive and with "(-)" if the association with the outcome is negative. For breast cancer treatment, which was represented by three indicator variables as predictors, no attempt is made in the table to reflect the ordering of the group-specific results (which can be discerned from subsequent tables), and the significance finding refers to the result of an omnibus significance test.

The second column of the table similarly shows results from fitting single-variable models to predict the number of reported problems on the domain using a truncated negative binomial model. The third column shows results from simple linear regressions using the sum of PAOFI values associated with problem items (where reverse coding was done), and the fourth column shows results from simple linear regressions using the average of PAOFI values associated with problem items.

It is noteworthy that being unemployed appears in all of the models, with the expected direction of unemployment associated with worse outcomes. Verbal memory neuropsychological test also is a predictor in all models, as might be expected for Memory: Forgetfulness, with higher scores on the neuropsychological test associated with lower Memory: Forgetfulness scoress (i.e., better outcomes). Furthermore, a number of predictors have significant associations with the outcomes in 3 of the 4 models (breast cancer treatment group, cancer stage at diagnosis, months since surgery, income<$100,000, education level, and BDI), while some covariates only appear in a single model (e.g., being married). An overarching conclusion is that different covariates are predictive of different aspects of forgetfulness, reinforcing the motivation for parsing PAOFI outcomes in this manner.

Similarly, Table 5.2 shows results for the Memory: Absent-Mindedness domain, Table 5.3 shows results for the Language domain, Table 5.4 shows results for the Sensorimotor domain, and Table 5.5 shows results for the Higher-Level Cognition domain. As might be expected, certain predictors appear frequently in the tables, but there remain differences

in columns reflecting existence of different predictors for distinct aspects of quality-of-life issues. We found breast cancer treatment group to be an important predictor for both memory domains, and the Language domain, as it appeared significant in predicting any issues versus none, the number of issues, and the cumulative symptom burden for each of these domains. Breast cancer treatment group was a significant predictor of having any issues with higher-level cognition, but not of any other aspects of the domain. It was not significant for any models fit to the Sensorimotor domain.

5.3 Three-part models with multiple predictors

A natural next step after considering candidate predictors one at a time is to consider multiple predictors of the different aspects (existence of problems, number of problems, severity of problems) of quality-of-life issues conveyed by responses to PAOFI questions. Because of the vast array of potential multivariable analyses that could be considered, the approach taken here was to organize analyses based on groups of predictors reflecting similar types of content. The predictors age, non-white race/ethnicity, married, education level, unemployed, and income $>$ \$100,000 were viewed as reflecting demographic and economic characteristics; the predictors breast cancer treatment group, breast cancer stage at diagnosis, months since surgery, months since last treatment, hormone replacement therapy, and body mass index were viewed as reflecting clinical characteristics. The predictors Beck Depression Inventory (BDI), quality-of-life (QOL), and Wechsler Test of Adult Reading (WTAR) were taken as reflecting psychological characteristics. And the predictors verbal memory, verbal learning, visual memory, visual learning, visual-spatial, executive function, motorspeed, and psychmotor were viewed as reflecting neuropsychological test performance.

Within a given factor domain, with the corresponding ordinal-valued PAOFI items contributing to a composite variable treated as the outcome in the various analyses investigated here (reversing the ordinal scale when indicated so that higher scores imply worse quality

of life), we started with a model including only breast cancer treatment group indicators. We also carried out analyses that forced the inclusion of breast cancer treatment group indictors in all models while also exploring (using forward variable selection) an additional group of predictor variables, where the group of predictor variables was either the demographic/economic, clinical, psychological, or neuropsychological-test-performance group of variables. Covariates were added sequentially based on satisfying criterion of $p < .1$.

Table 5.6 illustrates the approach using the composite score for the Memory: Absent-Mindedness factor as an outcome and reporting coefficients from each of five fitted logistic-regression models for existence of any problem, including the model with only breast-cancer-treatment-group indicators and each of four other models with an additional group of covariates considered. Not all of the models gave rise to significant predictors after controlling for treatment group, so the findings for those fitted models are the same as for the model with only treatment-group indicators. As might have been anticipated, there are instances where additional predictor variables emerge as significant after controlling for treatment-group indicators. For example, after accounting for the effect of treatment, BDI ($p < .05$) and QOL ($p < .05$) were significant predictors of experiencing absent-mindedness, with signs in opposite directions, underscoring the need to consider specific covariate values to accurately predict probability of experiencing a problem. In the model with demographic covariates, unemployment was a significant predictor ($p < .05$) with coefficient corresponding to odds ratio of $e^{.71} = 2.03$. Thus, for a white patient who received chemo and radiation therapy, the predicted probability of any absent-mindedness would be approximately 52 percent if the person were employed, and nearly 69 percent if the patient were unemployed.

In similar fashion, Table 5.7 reports findings for the Memory: Absent-Mindedness factor using treatment-group and blocks of other predictors in truncated-negative-binomial models for the number of reported problems. Controlling for the effect of breast cancer treatment, more issues with absent-mindedness were associated with BMI ($p < .05$) and BDI ($p < .05$). For example, the expected number of symptoms from truncated negative binomial for a

patient with chemo and radiation therapy, and a BMI of 22 (normal weight) would be $e^{-1.36+.3+.05*22} = 1.04$. The estimated number of symptoms would be 1.34 and 1.72 for chemo and radiation patients with BMI=27 (in mid-range of overweight) and BMI=32 (borderline obese), respectively. In the model with psychological covariates, a patient with chemo and radiation is expected to report approximately 1.40 symptoms if she has BDI=10 (minimum range); 1.8 symptoms if her BDI=18 (mild range); 2.27 symptoms if her BDI=26 (moderate range); and 2.89 symptoms if her BDI=34 (severe range).

Table 5.8 reports findings for the Memory: Absent-Mindedness factor in linear regression models on log scale for the aggregate (sum) severity score. Controlling for the effect of breast cancer treatment group, a 10-point increase in BMI was associated with a 22 percent increase in the cumulative symptom burden score ($p < .1$).

Table 5.9 reports analogous findings for the Memory: Absent-Mindedness factor in linear regression models for the average severity score. Controlling for the effect of breast cancer treatment group, being unemployed was associated with an additional .25 in the average symptom severity score ($p < .05$).

5.4 Discussion

We proposed a multi-part model that improves upon the widely-used dichotomization of the PAOFI by complementing such a dichotomization with additional analyses of subparts of the data. Specifically, we suggest logistic model for any symptoms versus none, a count model for the number of experienced symptoms, and a linear regression for the average severity, and/or the cumulative burden due to symptoms. As anticipated, different predictors emerge as significant in different models, consistent with the motivation for parsing PAOFI outcomes into different components.

In addition to each PAOFI item being measured on an ordinal scale, where theory would suggest that dichotomization would result in loss of information, the fact that the scale

includes 33 items asking about different dimensions of quality of life gives rise to doubt that a one-number summary would capture all of the elements of quality of life that are important for breast cancer patients. The role of different predictors for distinct outcomes was seen both across unvariable analyses and multivariable analyses. That said, after controlling for breast cancer treatment group, we did not find a large number of additional predictors. For predicting the number of problems in a given domain and either the average severity or aggregate impact of problems in a given domain, the finding of there not being many significant predictors might be related to limitations in sample size due to only those patients who reported issues being included in the model.

Table 5.1: Significant predictors of difficulties with Memory: Forgetfulness from univariate analyses

Outcome			
Any issues vs none	Number of issues	Cumulative symptom burden	Average severity of issues
BC treatment group	*BC treatment group*	*BC treatment group*	***Hormone replacement therapy (+)***
Months since surgery (+)	***Months since surgery (+)***	***Months since surgery (+)***	Months since last Tx (+)
BC stage at Dx (-)	BC stage at Dx (+)	***Unemployed (+)***	BC stage at Dx (-)
Unemployed (+)	***Unemployed (+)***	*Income < $100,000 (+)*	Age (+)
Income < $100,000 (+)	*Income < $100,000 (+)*	Education level (+)	***Married (-)***
Education level (-)	*Education level (-)*	***BDI (+)***	Unemployed (+)
BDI (+)	***BDI (+)***	QOL (-)	***Verbal memory NP test (-)***
QOL (-)	***Verbal memory NP test (-)***	***Verbal memory NP test (-)***	***Visual memory NP test (+)***
WTAR (-)	Verbal learning NP test (-)	Visual memory NP test (+)	
Verbal memory NP test (-)	Visual memory NP test (+)	*Motorspeed NP test (-)*	
Executive function NP test (-)	***Motorspeed NP test (-)***		
Psychmotor NP test (-)			

Notes:

Bold italic indicates significance at .05 level; italic indicates significance at .10 level; otherwise, significance is at .25 level.

"(+)" indicates positive association with the outcome; "(-)" indicates negative association with the outcome.

"BC": breast cancer

"Dx": diagnosis

"Tx": treatment

"NP": neuropsychological

"BDI": Beck Depression Inventory

"QOL": quality-of-life indicator

Table 5.2: Significant predictors of difficulties with Memory: Absent-Mindedness from univariate analyses

Outcome			
Any issues vs none	Number of issues	Cumulative symptom burden	Average severity of issues
BC treatment group	*BC treatment group*	***BC treatment group***	***Unemployed (+)***
Months since surgery (+)	***BMI (+)***	Months since surgery (+)	Executive function NP test (-)
BC stage at Dx (+)	*Unemployed (+)*	*BMI (+)*	Psychmotor NP test (-)
Non-white race/ethnicity (+)	**BDI (+)**	*Unemployed (+)*	
Unemployed (+)	QOL (-)	***BDI (+)***	
Income < $100,000 (+)	Verbal learning NP test (+)	*Verbal learning NP test (+)*	
Education level (+)	Visual memory NP test (+)	Visual memory NP test (+)	
BDI (+)	Executive function NP test (-)	Psychmotor NP test (-)	
QOL (-)	Psychmotor NP test (-)		
Verbal memory NP test (-)			
Executive function NP test (-)			
Psychmotor NP test (-)			

Notes:

Bold italic indicates significance at .05 level; italic indicates significance at .10 level; otherwise, significance is at .25 level.

"(+)" indicates positive association with the outcome; "(-)" indicates negative association with the outcome.

"BC": breast cancer

"Dx": diagnosis

"Tx": treatment

"NP": neuropsychological

"BDI": Beck Depression Inventory

"QOL": quality-of-life indicator

Table 5.3: Significant predictors of difficulties with Language from univariate analyses

Outcome			
Any issues vs none	Number of issues	Cumulative symptom burden	Average severity of issues
BC treatment group	BC treatment group	BC treatment group	*Hormone replacement therapy (+)*
Months since last Tx (-)	Months since surgery (+)	Months since last Tx (-)	Months since last Tx (-)
Months since surgery (+)	Income < $100,000 (-)	Months since surgery (+)	Married (-)
BC stage at Dx (+)	***Verbal memory NP test (+)***	*BC stage at Dx (+)*	***Unemployed (+)***
Non-white race/ethnicity (+)		BMI (-)	*Education level (-)*
BDI (+)		Married (+)	***WTAR (-)***
QOL (-)		Income < $100,000 (-)	*Verbal memory NP test (-)*
WTAR (-)		BDI (+)	Verbal learning NP test (-)
Visual memory NP test (+)		***Verbal memory NP test (+)***	Visual-spatial NP test (-)
Executive function NP test (-)		*Visual memory NP test (+)*	***Executive function NP test (-)***
Psychmotor NP test (-)		Motorspeed NP test (-)	Psychmotor NP test (-)

Notes:

Bold italic indicate significance at .05 level; italic indicates significance at .10 level; otherwise, significance is at .25 level.

"(+)" indicates positive association with the outcome; "(-)" indicates negative association with the outcome.

"BC": breast cancer

"Dx": diagnosis

"Tx": treatment

"NP": neuropsychological

"BDI": Beck Depression Inventory

"QOL": quality-of-life indicator

"WTAR": Wechsler Test of Adult Reading

Table 5.4: Significant predictors of difficulties with Sensorimotor function from univariate analyses

Outcome			
Any issues vs none	Number of issues	Cumulative symptom burden	Average severity of issues
Hormone replacement therapy (+)	BMI (-)	Unemployed (+)	Months since surgery (-)
BDI (+)	Married (+)	Visual learning NP test (-)	BMI (+)
QOL (-)	*Unemployed (+)*	Motorspeed NP test (-)	Married (-)
WTAR (-)	Education level (-)	***Psychmotor NP test (-)***	WTAR (+)
Executive function NP test (-)	Verbal memory NP test (-)		Motorspeed NP test (-)
	Verbal learning NP test (-)		
	Psychmotor NP test (-)		

Notes:

Bold italic indicates significance at .05 level; italic indicates significance at .10 level; otherwise, significance is at .25 level.

"(+)" indicates positive association with the outcome; "(-)" indicates negative association with the outcome.

"BC": breast cancer

"Dx": diagnosis

"Tx": treatment

"NP": neuropsychological

"BDI": Beck Depression Inventory

"QOL": quality-of-life indicator

"WTAR": Wechsler Test of Adult Reading

Table 5.5: Significant predictors of difficulties with Higher-Level Cognition from univariate analyses

Outcome			
Any issues vs none	Number of issues	Cumulative symptom burden	Average severity of issues
BC treatment group	***Hormone replacement therapy (+)***	***Hormone replacement therapy (+)***	Months since last Tx (-)
Hormone replacement therapy (+)	Age (+)	Unemployed (+)	***BDI (+)***
Months since surgery (+)	Unemployed (+)	Education level (+)	QOL (-)
Non-white race/ethnicity (+)	*Education level (-)*	***BDI (+)***	*Visual-spatial NP test (+)*
Unemployed (+)	*BDI (+)*	*QOL (-)*	Executive function NP test (-)
BDI (+)	QOL (-)	Visual memory NP test (+)	
QOL (-)	Verbal learning NP test (-)	*Visual learning NP test (+)*	
WTAR (-)	Visual learning NP test (+)	***Visual-spatial NP test (+)***	
Verbal memory NP test (-)	***Visual-spatial NP test (+)***	*Motorspeed NP test (-)*	
Visual memory NP test (+)	*Motorspeed NP test (-)*	***Psychmotor NP test (-)***	
Motorspeed NP test (-)	Executive function NP test (-)		
Executive function NP test (-)	***Psychmotor NP test (-)***		
Psychmotor NP test (-)			

Notes:

Bold italic indicates significance at .05 level; italic indicates significance at .10 level; otherwise, significance is at .25 level.

"(+)" indicates positive association with the outcome; "(-)" indicates negative association with the outcome.

"BC": breast cancer

"Dx": diagnosis

"Tx": treatment

"NP": neuropsychological

"BDI": Beck Depression Inventory

"QOL": quality-of-life indicator

"WTAR": Wechsler Test of Adult Reading

Table 5.6: Predictors of any difficulty with Memory: Absent-Mindedness from multiple logistic regression model

Predictor	BC Tx group	BC Tx group + demographics	BC Tx group + clinical	BC Tx group + psychological	BC Tx group + NP tests
Intercept	***-.88***	*-.63*	***-.88***	.20	-.56
Radiation alone	-.06	.01	-.06	.25	-.06
Chemo alone	.26	.19	.26	-.16	.22
Chemo & radiation	***.70***	***.70***	***.70***	***.77***	***.68***
Non-white race/ethnicity		*.69*			
Unemployed		***.71***			
Beck Depression Inventory				***.13***	
Quality of life				***-.33***	
Verbal memory NP test					*-.47*

Notes:

Table entries refer to coefficients from multiple logistic regression model.

Bold italic indicates significance at .05 level; italic indicates significance at .10 level.

After controlling for BC treatment group, none of the clinical variables were significant at .10 level.

"BC": breast cancer

"Tx": treatment

"NP": neuropsychological

Table 5.7: Predictors of number of difficulties with Memory: Absent-Mindedness from multiple truncated negative binomial regression

Predictor	BC Tx group	BC Tx group + demographics	BC Tx group + clinical	BC Tx group + psychological	BC Tx group + NP tests
Intercept	***-.17***	***-.17***	*-1.36*	-.41	***-.17***
Radiation alone	-.11	-.11	-.39	-.01	-.11
Chemo alone	*1.25*	*1.25*	1.00	.89	*1.25*
Chemo & radiation	.61	.61	.30	.45	.61
Body Mass Index			***.05***		
Beck Depression Inventory				***.03***	

Notes:

Table entries refer to coefficients from multiple truncated negative binomial regression model.

Bold italic indicates significance at .05 level; italic indicates significance at .10 level.

After controlling for BC treatment group, none of the demographic and neuropsychological tests variables were significant at .10 level.

"BC": breast cancer

"Tx": treatment

"NP": neuropsychological

Table 5.8: Predictors of aggregate impact of difficulties with Memory: Absent-Mindedness from multiple linear regression

Predictor	BC Tx group	BC Tx group + demographics	BC Tx group + clinical	BC Tx group + psychological	BC Tx group + NP tests
Intercept	***1.76***	***1.76***	***1.25***	***1.76***	***1.76***
Radiation alone	-.11	-.11	-.20	-.11	-.11
Chemo alone	***.76***	***.76***	*.67*	***.76***	***.76***
Chemo & radiation	.21	.21	.09	.21	.21
Body Mass Index			*.02*		

Notes:

Table entries refer to coefficients from multiple linear regression model where the outcome is log (sum PAOFI score).

Bold italic indicates significance at .05 level; italic indicates significance at .10 level.

After controlling for BC treatment group, none of the demographic, psychological, and neuropsychological tests variables were significant at .10 level.

"BC": breast cancer

"Tx": treatment

"NP": neuropsychological

Table 5.9: Predictors of average severity of difficulties with Memory: Absent-Mindedness from multiple linear regression

Predictor	BC Tx group	BC Tx group + demographics	BC Tx group + clinical	BC Tx group + psychological	BC Tx group + NP tests
Intercept	***4.13***	***4.00***	***4.13***	***4.13***	***4.13***
Radiation alone	-.01	.01	-.01	-.01	-.01
Chemo alone	.32	.26	.32	.32	.32
Chemo & radiation	.11	.12	.11	.11	.11
Unemployed		***.25***			

Notes:

Table entries refer to coefficients from multiple linear regression model where the outcome is average PAOFI score.

Bold italic indicates significance at .05 level; italic indicates significance at .10 level.

After controlling for BC treatment group, none of the clinical, psychological, and neuropsychological tests variables were significant at .10 level.

"BC": breast cancer

"Tx": treatment

"NP": neuropsychological

Bibliography

Afifi, A., May, S., and Clark, V. (2011). *Practical Multivariate Analysis*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press.

Agresti, A. (2010). *Analysis of Ordinal Categorical Data*. Wiley Series in Probability and Statistics. Wiley.

Beckstead, J. W. and Beckie, T. M. (2011). How much information can metabolic syndrome provide? an application of information theory. *Medical Decision Making* 31, 79–92.

Belotti, F., Deb, P., Manning, W. G., Norton, E. C., et al. (2015). twopm: Two-part models. *Stata J* 15, 3–20.

Bender, C. M., Pacella, M. L., Sereika, S. M., Brufsky, A. M., Vogel, V. G., Rastogi, P., Casillo, F. E., Richey, S. M., and Ryan, C. M. (2008). What do perceived cognitive problems reflect? *The journal of supportive oncology* 6, 238.

Bender, C. M., Sereika, S. M., Berga, S. L., Vogel, V. G., Brufsky, A. M., Paraska, K. K., and Ryan, C. M. (2006). Cognitive impairment associated with adjuvant therapy in breast cancer. *Psycho-Oncology* 15, 422–430.

Bradlow, E. T. and Zaslavsky, A. M. (1999). A hierarchical latent variable model for ordinal data from a customer satisfaction survey with "no answer" responses. *Journal of the American Statistical Association* 94, 43–52.

Chelune, G. J., Heaton, R. K., and Lehman, R. A. (1986). Neuropsychological and personality correlates of patients' complaints of disability. In *Advances in clinical neuropsychology*, pages 95–126. Springer.

Cover, T. M. and Thomas, J. A. (1991). Entropy, relative entropy and mutual information. *Elements of Information Theory* 2, 1–55.

Dawson, N. V. and Weiss, R. (2012). Dichotomizing continuous variables in statistical analysis a practice to avoid. *Medical Decision Making* 32, 225–226.

Duan, N., Manning, W. G., Morris, C. N., and Newhouse, J. P. (1984). Choosing between the sample-selection model and the multi-part model. *Journal of Business & Economic Statistics* 2, 283–289.

Ganz, P. A., Bower, J., Kwan, L., Castellon, S., Silverman, D., Geist, C., Breen, E., Irwin, M., and Cole, S. (2013). Does tumor necrosis factor-alpha (tnf-α) play a role in post-chemotherapy cerebral dysfunction? *Brain, behavior, and immunity* 30, S99–S108.

Ganz, P. A., Kwan, L., Castellon, S. A., Oppenheim, A., Bower, J. E., Silverman, D. H. S., Cole, S. W., Irwin, M. R., Ancoli-Israel, S., and Belin, T. R. (2013). Cognitive complaints after breast cancer treatments: Examining the relationship with neuropsychological test performance. *Journal of the National Cancer Institute* 105, 791–801.

Hays, R. D., Morales, L. S., and Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical care* 38, II28.

Hunter, J. E. and Schmidt, F. L. (1990). Dichotomization of continuous variables: The implications for meta-analysis. *Journal of Applied Psychology* 75, 334.

Johnson, V. and Albert, J. (2013). *Ordinal Data Modeling*. Statistics for Social and Behavioral Sciences. Springer New York.

MacCallum, R. C., Zhang, S., Preacher, K. J., and Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological methods* 7, 19.

Pullens, M. J., De Vries, J., and Roukema, J. A. (2010). Subjective cognitive dysfunction in breast cancer patients: a systematic review. *Psycho-Oncology* 19, 1127–1138.

Ragland, D. R. (1992). Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. *Epidemiology* 3, 434–440.

Ribaudo, H. J., Thompson, S. G., and Allen-Mersh, T. G. (2000). A joint analysis of quality of life and survival using a random effect selection model. *Statistics in Medicine* 19, 3237–3250.

Richardson-Vejlgaard, R., Dawes, S., Heaton, R. K., and Bell, M. D. (2009). Validity of cognitive complaints in substance-abusing patients and non-clinical controls: The patient's assessment of own functioning inventory (paofi). *Psychiatry research* 169, 70–74.

Rourke, S. B., Halman, M. H., and Bassel, C. (1999a). Neurocognitive complaints in hiv-infection and their relationship to depressive symptoms and neuropsychological functioning. *Journal of Clinical and Experimental Neuropsychology* 21, 737–756. PMID: 10649531.

Rourke, S. B., Halman, M. H., and Bassel, C. (1999b). Neuropsychiatric correlates of memory-metamemory dissociations in hiv-infection. *Journal of Clinical and Experimental Neuropsychology* 21, 757–768. PMID: 10649532.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal* 27, 379–423, 623–656.

Shelton, M. D. and Parsons, O. A. (1987). Alcoholics' self-assessment of their neuropsychological functioning in everyday life. *Journal of Clinical Psychology* 43, 395–403.

Song, M.-K., Ward, S. E., Bair, E., Weiner, L. J., Bridgman, J. C., Hladik, G. A., and Gilet, C. A. (2015). Patient-reported cognitive functioning and daily functioning in chronic dialysis patients. *Hemodialysis International* 19, 90–99.

Strömberg, U. (1996). Collapsing ordered outcome categories: a note of concern. *American journal of epidemiology* 144, 421–424.

Suissa, S. (1991). Binary methods for continuous outcomes: a parametric alternative. *Journal of clinical epidemiology* 44, 241–248.

Sy, J. P. and Taylor, J. M. (2000). Estimation in a cox proportional hazards cure model. *Biometrics* 56, 227–236.

Taylor, A. B., West, S. G., and Aiken, L. S. (2006). Loss of power in logistic, ordinal logistic, and probit regression when an outcome variable is coarsely categorized. *Educational and psychological measurement* 66, 228–239.

Van Dyk, K., Ganz, P. A., Ercoli, L., Petersen, L., and Crespi, C. M. (2016). Measuring cognitive complaints in breast cancer survivors: psychometric properties of the patient's assessment of own functioning inventory. *Supportive Care in Cancer* 24, 4939–4949.