

Spring 2012

Should we change the way we model change? Comparing traditional and modern techniques in modeling change in sense of identity over time

Kelli Marie Samonte
James Madison University

Follow this and additional works at: <https://commons.lib.jmu.edu/master201019>

 Part of the [Psychology Commons](#)

Recommended Citation

Samonte, Kelli Marie, "Should we change the way we model change? Comparing traditional and modern techniques in modeling change in sense of identity over time" (2012). *Masters Theses*. 311.
<https://commons.lib.jmu.edu/master201019/311>

This Thesis is brought to you for free and open access by the The Graduate School at JMU Scholarly Commons. It has been accepted for inclusion in Masters Theses by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.

Should We Change the Way We Model Change? Comparing Traditional and Modern
Techniques in Modeling Change in Sense of Identity Over Time

Kelli Samonte

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Department of Graduate Psychology

May 2012

Acknowledgements

I would like to acknowledge the individuals who have supported me through this process. First, I would like to thank my advisor and committee chair, Dr. Dena Pastor, for her undying support and dedication. Dena's encouragement, feedback, and guidance throughout my entire graduate career have enabled me to accomplish more than I could have ever imagined. Her guidance has truly helped me get closer to becoming the person I hope to be. I can never thank her enough for all that she has taught me both academically and professionally.

In addition, I would like to thank my committee members, Dr. Sara Finney and Dr. Chris Hulleman, for the time they took to read my thesis and provide guidance. Their thorough, thoughtful feedback and innovative ideas greatly contributed to the quality of this document. Moreover, their questions often led me to delve deeper and expand my understanding of the literature and the techniques explored in this study. I am exceedingly grateful for all of their wonderful contributions to this project.

I would also like to thank my fellow graduate students for all of their support. My colleagues were always there to encourage me through the difficult issues, provide opinions and guidance, and take my mind off the subject with relaxing stir fry dinners. Last but not least, I would like to thank my family, friends, and partner for their unceasing patience and love. No matter how engrossed I am with academics, I know that they will always be there to support, encourage, and love me. I am incredibly lucky to have such wonderful individuals in my life.

Table of Contents

Acknowledgements.....	ii
List of Tables	v
List of Figures.....	vi
Abstract.....	vii
CHAPTER I: Introduction	1
Introduction to Techniques for Analyzing Longitudinal Data.....	1
Residual Covariance Matrices Used in Longitudinal Data Analysis.....	2
Techniques and Covariance Matrices	7
Types of Data.....	9
Techniques and Types of Data.....	11
Focus of Techniques	13
Summary.....	14
Overview.....	16
CHAPTER IIA: Review of the Literature	18
Theoretical Conceptualizations of Identity Throughout History	18
Measurement of Identity.....	22
Sense of Identity Scale.....	27
Correlates/Importance of Identity.....	29
Identity in Higher Education.....	30
Growth/Change Over Time of Identity.....	33
Longitudinal Research in Identity.....	34
CHAPTER IIB: Residuals	36
Traditional Regression.....	36
Regression with a Categorical Predictor.....	40
Repeated Measures Data.....	46
ACS Modeling	54
Choosing Among Covariance Matrices	57
Limitations of ACS Modeling	58
Multilevel Models for Longitudinal Data.....	60
Unconditional means model.....	62
Unconditional growth model.....	64
CHAPTER III: Method.....	70

Overview of Analyses.....	70
Participants and Procedure.....	73
Phase 1.....	73
Phase 2.....	74
Measure.....	75
Part A.....	75
Part B.....	77
Part C.....	78
CHAPTER IV: Results.....	83
Part A.....	83
Comparing PROC GLM and PROC MIXED.....	83
Applied Example.....	84
Part B.....	87
ACS Models.....	87
Multilevel Models.....	92
Overall Comments.....	96
Part C.....	97
CHAPTER V: Discussion.....	104
Limitations and Future Directions.....	112
Final Conclusions.....	114
Appendix A.....	115
Appendix B.....	117
Appendix C.....	119
Appendix D.....	121
Appendix E.....	124
References.....	125

List of Tables

- Table 1: First 12 Example Individuals for Traditional Regression
- Table 2: Predicted Scores and Residuals for First 12 Example Individuals
- Table 3: Two Dummy Codes to Represent Categorical Variable t
- Table 4: Three Dummy Codes to Represent Categorical Variable t
- Table 5: Predicted Scores and Residuals for First 12 Example Individuals with Categorical Predictor
- Table 6: Demonstration of the Person Effect
- Table 7: Residual Covariance Structures
- Table 8: Summary of Sense of Identity Data Collection
- Table 9: Descriptive Statistics for the Type I Dataset
- Table 10: Comparison of PROC GLM and PROC MIXED
- Table 11: Comparing Models
- Table 12: Comparing Random Effects Parameters for ACS Models
- Table 13: Comparing Fixed Effects Parameters
- Table 14: Comparing Random Effects Parameters for Multilevel Models
- Table 15: Fixed and Random Effects for the Unconditional Mean and Unconditional Growth Models
- Table 16: Descriptive Statistics for Time Variable

List of Figures

- Figure 1: James Marcia's four categories of identity development
- Figure 2: Residual values in traditional regression
- Figure 3: Matrix notation for the assumption of normally distributed residuals
- Figure 4: Residual values in regression with a categorical predictor
- Figure 5: Demonstration of the assumption of independent observations
- Figure 6: Residual values in repeated measures models
- Figure 7: Formation of the \mathbf{V} matrix from the \mathbf{G} and \mathbf{R} matrices
- Figure 8: Combination of \mathbf{G} and \mathbf{R} to form \mathbf{V} in the random intercept, random slope model
- Figure 9: Measurement schedule and trajectories for 25 participants

Abstract

Repeated measures analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA) are two of the most common techniques employed in longitudinal data analysis. These methods, however, are extremely limited in the type of data permitted in analysis, the residual covariance matrices employed in analysis, as well as in the focus of the research questions. There are, however, modern techniques for analyzing longitudinal data that do not have the same limitations of repeated measures ANOVA and MANOVA. This study aims to compare traditional methods of analyzing longitudinal data with more modern techniques, including alternative covariance structure (ACS) modeling and multilevel modeling (MLM), through an example involving Sense of Identity in college students. This is done by first exploring assumptions of traditional and modern methods of analyzing longitudinal data. Next, an introduction to the identity literature is provided. The concept of residuals in between- and within-subjects analyses is then discussed. Finally, both traditional and modern techniques are employed to analyze the Sense of Identity data and results are compared and contrasted in an attempt to demonstrate the utility and benefits of more advanced techniques in longitudinal data analysis.

CHAPTER I: Introduction

Introduction to Techniques for Analyzing Longitudinal Data

There are several methods used in practice to analyze longitudinal data, some being more commonly utilized than others. The purpose of this chapter is to provide a summary of a few possible techniques as well as a rationale as to which techniques would be more appropriate than others, depending on the assumptions, situation, and research questions at hand.

Some of the more traditional techniques used for analyzing longitudinal data include procedures like repeated measures analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA). Both repeated measures ANOVA and MANOVA are taught in introductory and intermediate statistics courses and are fairly easy to employ with common statistical software packages. Because of their familiarity and simplicity, it is no surprise that these models are commonly used to examine longitudinal data. In exchange for this familiarity and simplicity, however, these models make strict assumptions about the type of data and the structure of the residual covariance matrix, as will be explained in detail later. A more modern technique used to analyze longitudinal data that may be less well known than repeated measures ANOVA and MANOVA is Alternative Covariance Structure (ACS) modeling using PROC MIXED in SAS (SAS Institute Inc., 1992). ACS modeling is mildly more complex than the traditional techniques but offers some advantages including the type of data that can be used and the residual covariance matrices that can be applied.

Another modern technique that can be used to analyze longitudinal data and offers several advantages over the more traditional techniques is multilevel modeling.

Multilevel modeling (MLM), also known as hierarchical linear modeling (HLM), is a regression technique used with nested data, or data in which the assumption of independence of observations is violated. In longitudinal analyses, an individual is measured at several time points. Thus, the data is nested in that measurement occasions are nested within people. It would be inappropriate to assume that one observation from one individual would be independent from another observation from the same individual. Multilevel modeling yields advantages similar to ACS modeling, with the additional benefit of examining individual differences in change over time. The other techniques to analyze longitudinal data listed thus far focus primarily on overall change and not individual variability in change, whereas multilevel modeling allows for examination of both.

In the sections that follow, the role of residual covariance matrices in longitudinal data analysis is discussed and some of the possible structures for residual covariance matrices are presented. Only a brief treatment of residuals and residual covariance matrices is provided here as a more thorough treatment of these topics is provided in Chapter IIB. This section is followed by an overview of the types of longitudinal data that can be collected as well as which techniques can be used with particular types of data. Additionally, the advantages and disadvantages of the techniques with respect to their assumptions about residual covariance matrices and types of data with which they can be employed are provided.

Residual Covariance Matrices Used in Longitudinal Data Analysis

All of the techniques mentioned above investigate mean differences in the dependent variable across time points. In order to come to accurate conclusions regarding

these mean differences, however, data must satisfy assumptions made about 1) the variability of the residuals at different time points and 2) how the residuals covary between time points. “Residuals” can be defined as the difference between individuals’ observed scores and their respective predicted score based on the model specified. Residual variances and covariances provide information about the spread of scores at levels of the independent variable and about the relationship of the scores between different levels of the independent variable, respectively. There are several different formats that are possible for the residual variances and covariances resulting in different covariance structures. The following overview is by no means an exhaustive list of the possible covariance structures but demonstrates the similarities and differences among a few of the possibilities.

In order to describe each of the covariance structures, an example in which students have responded to a scale that measures some construct, Y , at three time points will be used. The first covariance structure to consider is the *compound symmetry residual covariance matrix*. In this matrix, the residual variances for time points 1, 2, and 3, which represent the spread of scores, are set equal, meaning that the variability in the residuals at all three time points is exactly the same. The residual covariances between all of the time points are also set to be equal, indicating that the covariances between the residual scores at time points 1 and 2, 1 and 3, and 2 and 3 are equivalent. Compound symmetry is appealing because only two parameters (one variance and one covariance) need to be estimated. That being said, suggesting that every time point has the same residual variance and that the relationship between all of the time points is the same is an incredibly strict assumption and may be considered overly restrictive in some situations.

For example, one might believe that in the example individuals who are more mature may be less variable in Y than less mature individuals. In this situation, we would expect the residual variance of Y at time one to be much larger than the residual variance of Y at time three. It also may be plausible that the relationship between residuals at adjacent time points (time 1 and time 2 or time 2 and time 3) would be stronger (larger) than between residuals at time points that are nonadjacent (time 1 and time 3) (Singer & Willett, 2003). If models that assume a compound symmetric residual covariance structure are specified for data that violates the compound symmetric assumption, standard errors can be biased. In this sense it would be appealing to apply a technique that allows the residual variances and covariances to be freely estimated.

An *unstructured residual covariance matrix* allows for just that situation. In an unstructured residual covariance matrix each residual variance and each residual covariance is freely estimated. Thus, the residual variance for Y at time 1, 2, and 3 can be three different values. Conceptually, different residual variances across time would suggest that the spread of scores differs across levels of the independent variable (time in repeated measures data). Thus, individuals' scores are more alike (smaller residual variance) or more different (larger variance) for different measurement occasions. The same goes for residual covariances in that the residual covariances between the residuals at time 1 and 2, time 2 and 3, and time 1 and 3 are all free to be whatever value the data suggests. The unstructured matrix is appealing because it is incredibly flexible in that every parameter (i.e., residual variances and covariances) can be freely estimated. Thus, because the unstructured residual covariance structure doesn't make assumptions about the residual variances and covariances, it allows for unbiased standard errors due to

violation of residual covariance structure assumptions. Its limitation, however, is that because all parameters are freely estimated it may be difficult to produce a precise solution if the sample size is not large enough. A precise solution is one in which the estimated parameters are stable and are not overly influenced by sampling error. In order to obtain precise estimates for each parameter, it is imperative that there are an adequate number of observations. As the number of parameters increases, the number of observations necessary to obtain precise estimates increases as well. Thus, freely estimating every residual variance and covariance can be especially problematic as the number of measurement occasions, and consequently the number of variances and covariances, increases. In this sense, this residual covariance matrix may seem overly complex and could possibly exploit idiosyncrasies in the data. In other words, the model could be over fitted to the data making it difficult to generalize to other samples. In our example, 6 parameters (three residual variances and three residual covariances) must be estimated as opposed to the 2 that needed to be estimated with the compound symmetry residual covariance matrix. Notably, as the number of time points increases, the number of parameters also increases. Thus, if researchers doubled the number of measurement occasions from 3 time points to 6 time points, the number of parameters estimated would increase from 6 to 21 (6 variances, and 15 covariances). As previously stated, a large number of parameters necessitate a large sample in order to obtain precise estimates. In sum, the compound symmetry structure makes incredibly strict assumptions about our residual variances and covariances, whereas the unstructured structure makes no assumptions about residual variances and covariances but may have difficulty acquiring

precise parameter estimates for the residual variances and covariances that are not too sample-specific. Where is the happy medium?

There are several residual covariance structures that fall into this “happy medium” category that will be described in detail in Chapter IIB. As an example, consider the *homogeneous autoregressive residual covariance structure*. Here, the residual variances are equal across time points; in other words, it assumes that the variability of the residual scores is exactly the same at each measurement occasion. Thus, the spread of scores is the same at each measurement occasion. If, for example, variances increased over time, it would suggest that scores are more spread out as time goes on. It also assumes that adjacent time points will have larger residual covariances than nonadjacent time points. Thus, the residuals for adjacent time points (e.g., time 1 and time 2) are more alike than the residuals for non-adjacent time points (e.g., time 1 and time 3). Measurement occasions that are temporally closer are often thought to be more alike than those that are further apart (Singer & Willett, 2003). In order to model these residual covariances, the ρ parameter, which captures the relationship between adjacent time points, is estimated. Because ρ represents the correlation between adjacent time points, it ranges from -1.0 to 1.0. The covariances are then expressed as a function of ρ and the variance (σ^2) in that the covariance between adjacent time points (e.g., time 1 and time 2), which are one step away from one another, are estimated as $\sigma^2\rho$. The covariances between time points that are two steps away (e.g., time 1 and time 3) are $\sigma^2\rho^2$, and so on. Although there are some restrictions as to the equality of variances and how the residual scores covary between time points, it is undoubtedly more flexible than the compound symmetric specification. This residual covariance structure is also appealing in terms of parsimony in that only

two parameters need to be estimated: σ^2 and ρ . Because measurement occasions that are temporally closer are often considered to be more related than those temporally farther apart, this residual covariance matrix is often considered in longitudinal research.

Techniques and Covariance Matrices

Different techniques for analyzing longitudinal data make different assumptions about the residual covariance matrices for the data. It is imperative that researchers analyzing longitudinal data consider the assumptions each technique makes about the residual covariance matrices and whether or not they align with what theory and empirical evidence would suggest about the residual variances and covariances. Specifically, it is imperative that researchers note when the assumptions do not align with empirical evidence or what theory would dictate because a disconnect between assumptions and theory may affect inferential tests of mean differences. Repeated measures ANOVA assumes a compound symmetric residual covariance matrix. The compound symmetric matrix, as discussed above, has an incredibly restrictive form. In actuality, a similar but less restrictive assumption known as “sphericity” is used and accepted in practice. Notably, as long as the assumption of sphericity can be satisfied, the inferential tests regarding mean differences will not be biased (Hoffman, in preparation). The assumption of sphericity assumes that the variances of the difference scores between time points are equal (Field, 2009). This assumption differs from compound symmetry, in which variances are assumed to be equal, in that sphericity allows for variances to differ across time points so long as the residual variance of the difference scores is equivalent. It is important to clarify that sphericity is a necessary but not sufficient condition that must be met in order to satisfy the compound symmetric

assumption. Because the assumption of compound symmetry is so difficult to satisfy, the acceptance of sphericity as an adequate condition allows for the traditional repeated measures ANOVA to be used in common practice. It should be noted that if the assumption of sphericity is not satisfied, the omnibus F test is too liberal, thus increasing the risk of Type I error. However, adjustments to the repeated measures ANOVA can be used to help account for violations of sphericity. The Huynh-Feldt and Greenhouse-Geisser corrections can be used to adjust the degrees of freedom by the extent to which sphericity has been violated, which is captured in an index known as epsilon. These corrections adjust the degrees of freedom based on an estimate of epsilon to make the omnibus F test more conservative (Hoffman, in preparation).

MANOVA assumes an unstructured residual covariance matrix. This matrix requires every parameter to be estimated and thus provides the optimum amount of information about the data. Because every parameter is estimated, however, MANOVA may have issues acquiring precise estimates for parameters as well as issues with capitalizing on idiosyncrasies in the data. In addition, the degrees of freedom used for the denominator of the F-statistic are based on the number of persons, not the number of total observations (each individual has multiple observations). Thus, the denominator degrees of freedom are smaller than repeated measures ANOVA and Type II errors may increase (Hoffman, in preparation).

ACS modeling and multilevel models allow for more variety in the kinds of residual covariance matrices that can be modeled. These techniques are more flexible in that they can model a residual covariance matrix deemed both parsimonious and appropriate based on what theory dictates and empirical evidence supports, rather than

what is assumed by the statistical technique. These methods have the capability to model compound symmetric and unstructured covariance matrices if the researcher considers them to be most appropriate for the circumstances. However, there are also several “happy medium” matrices (such as homogeneous autoregressive structure discussed above) that allow for a more customized, flexible residual covariance matrix that more adequately reflects the underlying theory and/or the empirical data. Multiple models can be fit to the data with ACS and multilevel modeling, each with a unique residual covariance matrix. The fit of the models with different residual covariance matrices can then be compared to one another using information criteria (e.g., Akaike Information Criterion, Bayesian Information Criterion). Models with nested residual covariance structures can be compared using the likelihood ratio test as well. The goal of testing several alternative models is to find the most parsimonious model that yields acceptable fit to the data. This flexibility in the structure of the residual covariance matrix makes ACS modeling and multilevel modeling more appealing and often more appropriate options in analyzing longitudinal data.

Types of Data

In addition to assumptions about the residual covariance matrix, statistical techniques also make assumptions about the type of longitudinal data that can be analyzed. There are three types of data that can be collected over multiple time points. Each type of longitudinal data can be described by schedules and waves (Singer & Willett, 2003). The data collection *schedule* indicates whether or not data was collected for participants at the same time points (with the same length of time for each individual between time points). In order for individuals to have the same schedule of data

collection, it is not necessary that the time between each measurement occasion is equal, only that the time between measurement occasions is the same across participants. An example in which participants would have the same schedules of data collection would be one in which one group testing session was administered at the beginning of the semester, one three weeks into the semester, and one at the end of the semester. Thus, each respondent participates in an initial measurement, one three weeks later, and one when the semester ends. An example in which individuals would not have the same schedule would be if participants were sent a survey three times throughout the semester and asked to respond at their leisure. In this case some participants would respond immediately whereas others may wait several weeks to respond. Thus, because each individual would have a different interval of time between responses, they would not have the same schedule.

The number of *waves* corresponds to how many times data from each individual was obtained (all of the time points, or only some of the time points). For example, using the semester example above, a student who responded to the test at all three time points would have three complete waves of data. Another participant may have only responded to two of the time points (they ignored an email or were absent on the testing day), in which case they would only have two complete waves of data.

As previously stated, each of the three types of data can be described using different combinations of schedules and waves (Wu, West & Taylor, 2009). *Type I* data is data that is balanced on time with complete data. Data that is “balanced on time” is data that is collected for all participants on the same schedule. “Complete data” is data in which each participant has the same number of waves of data collection. This type of data

is very difficult to collect because it requires that data from all participants are collected on the exact same schedule and that there is absolutely no missing data (which is incredibly unrealistic).

Type II data is data that is balanced on time but allows for missing data. Again, data that is balanced on time indicates that the schedule for collecting data was the same for each participant. The allowance of missing data indicates that not all participants supplied data on all waves. Type II data is more likely than Type I data because each individual does not need to have completed every single wave. It does, however, require a very strict schedule of data collection which can be logistically difficult to implement.

Type III data is data that is unbalanced on time and allows for missing data. Thus, each individual can have a different interval of time between their waves of data collection. In addition, participants can have data for any number of waves of data. Type III is fairly easy to gather because participants can give data whenever and however many times is possible. Type III data is, undoubtedly, the most flexible type of longitudinal data.

Techniques and Types of Data

As with the residual covariance structures, different techniques for analyzing longitudinal data also require different types of data. Both repeated measures ANOVA and MANOVA require Type I data. Because Type I data is all but impossible to obtain in reality, researchers often begin with a Type II data set and then use listwise deletion to handle missing data. Listwise deletion involves deleting participants or observations with any missing data. Listwise deleting missing data may give the researcher a “Type I” dataset, but the use of listwise deletion makes the strict assumption that participants are

not missing waves of data due to some systematic cause (e.g., data is missing completely at random). If this assumption is not satisfied, listwise deletion may lead to biased results (Schafer & Graham, 2002). The omission of individuals with missing data also depletes sample size and, in turn, reduces power.

ACS modeling offers some relief from the strict data assumptions placed on repeated measures ANOVA and MANOVA in that it allows for Type II data. This type of data still requires that the same schedule of data collection is used for all participants, but allows for missing data. Type II data is able to be used with this technique because ACS modeling uses maximum likelihood estimation and therefore all cases, even those with missing data, provide information used in parameter estimation (Enders, 2010).

Maximum likelihood estimation differs from ordinary least squares estimation (most often used with repeated measures ANOVA and MANOVA) in that it is an iterative process that produces parameter values for which the sample data are most likely to occur. Ordinary least squares, on the other hand, produces parameter values for which the prediction errors are a minimum. Additionally, the assumption made about why data is missing is less restrictive than the assumption in Type I data (that data is missing completely at random). Specifically, it assumes that missing responses are missing at random, meaning that the presence or absence of a response may be related to other variables in the data set, but not to the underlying value of that variable (Wu, West & Taylor, 2009; Schafer & Graham, 2002). Although the assumption that the data are missing at random is an untestable assumption, Schafer and Graham (2002) argue that the bias caused by typical violations of this assumption will not seriously bias parameter estimates when ML estimation is used. Thus, if researchers can systematically collect

data on the same schedule for all participants (e.g., scheduled test dates), ACS modeling is a very appealing option for analyzing longitudinal data.

Multilevel modeling offers even more flexibility in that it permits the use of Type III data. The fact that Type III data can be used with multilevel modeling is incredibly appealing because it provides researchers with flexibility in data collection and allows for the use of all data no matter what schedule was used or how many waves were collected. Allowing variation in schedules is convenient for researchers in that it requires much less planning and logistical work to make sure each individual has the exact same schedule. Multilevel modeling is also appealing when analyzing archival data, in which the researcher has no way of controlling data collection.

Focus of Techniques

In addition to assumptions about the residual covariance matrix and the type of data used with each technique, it is also important to consider the focus of each technique. The most notable difference between the focus of techniques for analyzing longitudinal data is that repeated measures ANOVA, MANOVA and ACS modeling all focus on change in the mean of scores over time (overall change), whereas multilevel modeling captures both changes in the mean scores over time as well as changes in individuals' scores over time (overall and individual change). In other words, repeated measures ANOVA, MANOVA, and ACS modeling provide information concerning overall change but little information, or information that is hard to interpret, to describe how individuals change over time. Multilevel modeling provides information about how persons, overall, start out on a construct and how they change on average over time, as well as whether persons start out at different levels of a construct and change

differentially over time. That is, with multilevel modeling, the focus broadens to include not only overall or average change, but the variability in how people change over time. For example, if, overall, there was no change in a variable over time, the traditional techniques and ACS modeling would imply that scores are stable across time. What if, however, some individuals increased on a construct whereas others decreased over time? The average trajectory across individuals may be stable which would imply no change, but in reality individuals are changing over time, just in different directions. Repeated measure ANOVA, MANOVA and ACS modeling would likely miss the information that individuals are changing in different directions and conclude that there is no change over time. Multilevel modeling, however, allows the researcher to examine both individual and overall change and thus would indicate that individuals vary greatly in how they change even though there appears to be no change overall.

Summary

In sum, all of the information about residual covariance matrices, types of data, and the focus of each technique should be used together to determine which method of analyzing longitudinal data would be most useful and appropriate in different situations. Repeated measures ANOVA and MANOVA both have the appealing qualities of being familiar, traditional techniques as well as being computationally simple. The familiarity and simplicity of repeated measures ANOVA is offset due to the strict assumptions placed on the type of data and residual covariance matrix. MANOVA assumes an extremely relaxed residual covariance matrix, but also requires the strictest form of data. Notably, the advantages associated with an unstructured covariance matrix are countered by the issues with precisely estimating numerous parameters and capitalizing on

idiosyncrasies in the data. It is also important to note that to estimate several parameters with precision one needs a large sample size which is extremely difficult to obtain, particularly when listwise deletion is simultaneously employed to satisfy the type of data assumption. Repeated measures ANOVA and MANOVA only differ in their residual covariance matrices and thus it is important to consider when each method is appropriate. If the assumption of sphericity is met, repeated measures ANOVA would provide accurate, parsimonious, and powerful results, whereas MANOVA would provide accurate, complex and (likely) under powered results. If the assumption of sphericity has been violated, repeated measures ANOVA may have biased standard errors (which will affect the inferential tests of mean differences), and thus MANOVA should be employed. Thus the traditional models used to analyze longitudinal data are not ideal unless the strict assumptions regarding the type of data and residual covariance matrices can be satisfied.

ACS modeling is less familiar than the traditional techniques, but offers other appealing properties. ACS modeling allows for a moderately less restrictive type of data as well as for a wide variety of residual covariance structures. The main limitation with ACS modeling is that collecting data in which participants all have the same schedule can be logistically demanding for researchers. Data that is balanced on time requires a lot of preparation at the front end of a study in addition to maintaining the specified schedule throughout the duration of the study. The data restriction also prevents longitudinal analysis on data that has already been collected, unless the data was collected with a schedule that was balanced on time.

Multilevel modeling is, undoubtedly the least restrictive technique to analyze longitudinal data in the assumptions regarding the types of data and residual covariance structure. Thus, it is ideal when residual covariance matrices are thought to deviate from compound symmetry, in addition to when data cannot, or is not, collected with a specific schedule. Multilevel modeling may be somewhat more computationally intensive, but the freedom gained with the type of data and residual covariance matrix is unique and worthwhile in comparison to the other techniques. In addition to the advantages of having less restrictive assumptions, multilevel modeling also allows for the examination of both overall and individual change over time. Thus, multilevel modeling provides richer information and can answer more complex research questions than the traditional models. In considering possible types of data, residual covariance structures, and the focus of different techniques, it is evident that multilevel modeling is unparalleled.

Overview

The purpose of this study is twofold. The first purpose of this study is to compare and contrast more traditional techniques (i.e., repeated measures ANOVA, MANOVA) and more modern techniques (i.e., ACS modeling, multilevel modeling) using an example with Sense of Identity data. Specifically, repeated measures ANOVA, MANOVA, ACS Modeling, and MLM are compared in terms of overall model fit, specific parameters, and substantive conclusions using data collected on a Sense of Identity scale from college students. Although the data has a Type III data structure, the data was altered to align with a Type I data structure for the purpose of comparing the four techniques. Second, the study aims to examine change in sense of identity over time. MLM models with the sense of identity data in its original Type III form were used to

examine overall change over time in sense of identity, as well as variability in individual intercepts and slopes.

The following chapter is divided into two parts to aid in explanation of two important areas. Chapter IIA provides an overview of the identity literature. Although ample research has been conducted in the field of identity, this chapter serves as a frame of reference for where our particular measure of sense of identity fits into the field. Chapter IIB provides a more thorough explanation of residuals than provided in the current chapter, beginning with traditional regression and progressing through residuals in repeated measures data. In addition, Chapter IIB discusses residual covariance matrices mentioned in the current chapter as well as several other residual covariance matrices in detail. Chapter IIB ends with an introduction to MLM and the traditional models used to examine change over time. Specifically the unconditional means model and two forms of the unconditional growth model are presented and explained.

CHAPTER IIA: Review of the Literature

Theoretical Conceptualizations of Identity Throughout History

The construct of identity has been an area of interest for many theorists as early as the late 1800's. Some of the most popular early work in identity theory was presented by Erik Erikson in his 1950 book *Childhood and Society*. In his book, Erikson presents eight stages of development that individuals must experience as they develop and mature. Within each stage, individuals must complete a task or resolve some crisis in order to move to the next stage. Failure to resolve one's crisis not only results in failure to progress to the next stage, but can also lead to negative consequences. Crisis in this sense is accepted as a crucial moment or turning point in an individual's life, as opposed to a threat of imminent disaster, as it may be more commonly conceptualized (Erikson, 1968). Each of Erikson's stages consists of criteria that an individual must meet through resolving his or her crisis before it is possible to move on to the successive stages of development.

The introduction of identity occurred in Erikson's fifth stage of development, termed "identity vs. role confusion," also referred to as identity achievement vs. identity diffusion. In this stage of development, adolescents begin to face tangible adult tasks and, "are now primarily concerned with what they appear to be in the eyes of others as compared with what they feel they are, and with the question of how to connect the roles and skills cultivated earlier with the occupational prototypes of the day" (Erikson, 1950, p. 261). Thus, the definition of the identity crisis includes both internal (who they are) and social-contextual (what they appear to be in the eyes of others) dimensions

(Schwartz, 2001), indicating that identity is as much an understanding of who one is internally as it is an understanding of who one is in different situations.

The ultimate goals in resolving one's identity crisis would be to develop one's unique identity as well as to avoid the negative consequences brought about by failing to resolve one's identity crisis. Erikson states that if an individual is stuck within the role confusion stage, delinquent and psychotic episodes are frequent. Also, in an attempt to avoid negative consequences and an unhealthy sense of self, individuals will over-identify with "heroes". Within the identity vs. role confusion stage Erikson postulates that in order for individuals to resolve their identity crisis, they must explore the possible choices for identity resolution and commit to the one that is most representative of their past selves and hopeful future selves. Erikson believed that through the exploration of possible selves and commitment to the most representative self, one would meet the criteria necessary to resolve and move on from the identity vs. role confusion stage. Completion of the identity vs. role confusion stage in development results in one's crystalized identity. According to Erikson's theory, resolving one's identity crisis marks the end of childhood and is necessary before moving on to the next stage in development, intimacy vs. isolation.

Over time, the construct of identity has grown and evolved, but many of the theories are still based on the basic concepts proposed by Erikson. Most notably, Erikson's theories influenced James Marcia's (1966) commonly used framework of identity status. Like Erikson, Marcia proposes that individuals define themselves in both internal and socio-contextual domains through a cycle of exploration and commitment. He refers to this process of exploration and commitment as a psychological task that

individuals must complete in order to form a crystallized identity. Thus, Marcia's "psychological task" corresponds with Erikson's identity crisis. Erikson and Marcia's conceptualizations of identity differ in that Marcia utilized Erikson's conceptualization of identity to form more detailed categorizations of identity development. According to Marcia, individuals can be characterized into four distinct categories of identity status based on the presence or absence of Erikson's two decision making components: exploration and commitment. Marcia defines exploration as the phase in which individuals explore and choose between possible alternative selves that are most representative as to how one could solve an issue or make a decision. Individuals in the exploration phase are actively exploring and considering viable possibilities in an attempt to choose an option that best represents themselves. Commitment indicates that an individual openly chooses and personally invests in an identity. This commitment may be due to the exploration of alternatives from a crisis period or due to goals derived externally, perhaps proposed by the individual's parents. Marcia combines the presence or absence of exploration and the presence or absence of commitment to form four distinct categories of identity status: identity achievement, moratorium, foreclosure, and identity diffusion (see Figure 1). Erikson's conceptualization of the identity crisis forms the two extremes of his identity status paradigm (identity achievement and identity diffusion), whereas the moratorium and foreclosure points are seen as somewhat intermediate points of identity status.

		Exploration	
		Present	Absent
Commitment	Present	Identity Achievement	Foreclosure
	Absent	Moratorium	Identity Diffusion

Figure 1. James Marcia's four categories of identity development

Identity achievement is characterized by individuals who have experienced a crisis phase and have both explored and committed to an identity on their own terms. Identity achievement is the ideal stage for adolescents in that the thorough exploration of alternatives and subsequent commitment to an option indicates that the individual has chosen an option that best exhibits their unique internal beliefs and values. Thus, the identity achievement stage is analogous to Erikson's crystallized identity. Identity diffusion is characterized by individuals who have or have not experienced the exploration phase and are distinguished by their lack of commitment. Individuals in the identity diffusion stage have not made a commitment and are completely uninterested in the thought of committing to one decision. Individuals in this stage are likely to abandon their current occupation or ideological stances if other desirable opportunities are presented with little to no hesitation. Thus, identity diffusion is similar to Erikson's conceptualization of role confusion. Individuals in the moratorium phase are distinguished because they are in the middle of the exploration phase. Moratorium individuals have not made a commitment, but are distinct from those in the identity diffusion phase in that moratorium individuals are actively exploring and considering

alternatives with the intent to make commitments. Foreclosure individuals are characterized by those individuals who have not experienced the exploration phase, but have expressed commitment. The lack of exploration suggests that these individuals are relying heavily on external influences (e.g., parental beliefs and values) to make decisions.

Measurement of Identity

In order to examine identity, how it changes over time, and make inferences about what is related to the formation of identity, instruments that measure identity and provide reliable and valid scores are necessary. Sound instruments to measure identity would be beneficial in several testing situations. Chickering (1999) highlights the idea that instruments to measure identity development would be particularly beneficial in higher education settings. Identity development instruments in higher education can be used to evaluate programs and interventions to provide insight as to what facilitates identity development. Due to the undeniable importance of the construct, several methods of measuring identity have been developed over the last few decades.

One of the first methods developed to measure identity was an interview format developed by James Marcia. In order to assign each individual to one of his four stages of identity development, Marcia used one on one interviews lasting between 15-30 minutes. In the interviews, individuals were evaluated on whether or not crisis and/or commitment were present in the domains of occupation, religion, and politics. Religion and politics were eventually combined into overall ideology. Interviews were recorded and then replayed, possibly several times, in order for raters to objectively evaluate individuals and subsequently place them into one of the four categories shown in Figure 1. Though

Marcia's interviews have supporting validity evidence, there are a few issues with this type of measurement. First, conducting one on one 15-30 minute interviews and then replaying them several times is not an efficient way to collect data, particularly if one is interested in collecting data from a large number of respondents.

Some may also take issue with the idea that individuals are being forced into four mutually exclusive categories. From a measurement standpoint, a generous amount of research has been conducted undermining the categorization of variables that can be considered continuous. MacCallum et al. (2002) indicates that dichotomizing, or categorizing, a continuous or "graduated" variable will result in a substantial loss of power and biased effect sizes. In order to argue that categorization is problematic in Marcia's paradigm, it is important to consider what continuum is being categorized. It is fair to argue that the four stages in Marcia's paradigm may not be what one would conventionally consider continuous. More specifically, identity diffusion and identity achievement are clearly at the extreme ends of identity development, but because foreclosure and moratorium do not have a set place along the continuum it would be difficult to argue for a set linear development through these stages.

The issue with categorization in Marcia's paradigm has to do with the dichotomization of the continuous exploration and commitment variables. One problem with this categorization is that no distinction can be made in exploration or commitment among individuals in the same category. For example, an individual placed in the moratorium category could be just entering their crisis and just starting to brainstorm alternatives without having done any exploration yet which would put them somewhere between diffusion and moratorium. On the other hand, they could be toward the end of

their exploration of alternatives and be getting ready to make a commitment which would put them at a more advanced level of exploration. Unfortunately, these same two individuals would be placed into the same exploration category of moratorium, preventing any distinction in their levels of exploration to be made. Thus an instrument that is able to measure identity on a continuous scale would allow for a less crude definition among individuals' identity status.

Given the issues with Marcia's interviews, it is worthwhile to consider the other methods that have been developed to gather information about identity. The Extended Version of the Objective Measure of Ego Identity Status (EOM-EIS-II; Bennion & Adams, 1986) and the Q Ego Identity Status (Q-EIS; Mallory, 1989) are based on Marcia's conceptualization and thus focus on the four categories of identity achievement, moratorium, foreclosure, and identity diffusion. The Erikson Psychosocial Inventory Scale (EPSI; Rosenthal, Gurney & Moore, 1981) seeks to measure whether the identity crisis, as a whole, has been resolved, while the Ego Identity Process Questionnaire (EIPQ; Balistreri, Busch-Rossnagel & Geisinger, 1995) seeks to measure individuals' scores on exploration and commitment.

Many of the instruments seem to have addressed the issue of simply categorizing individuals into mutually exclusive categories. Even though many of the instruments are based on Marcia's paradigm, many of them have some continuous measure within in all four categories or provide a continuous score for exploration and commitment, as opposed to forcing respondents into one mutually exclusive category without any knowledge of where in that category they lie. This allows researchers to compare

individuals within each stage as well as between stages and provides researchers with more information about the status of individuals' identity development.

A vast majority of the instruments used to measure identity formation throughout history also tended to capture individuals' identity within specific domains (e.g., race, gender, occupation, etc.). Erikson (1980) points out that identity and identity crisis in scientific research can be seen as constructs which, "circumscribe something so large and so seemingly self-evident that to demand a definition would almost seem petty, while at other times they designate something made so narrow for purposes of measurement that the over-all meaning is lost" (p.15). Erikson's acknowledgement of how broad the construct of identity is gives support for why researchers tend to break the concept of identity into specific domains. That being said, he also makes the point that by breaking the construct into more narrow, manageable pieces the true meaning of identity can get lost. In reality it would be fairly rare for one to think of themselves solely in terms of their occupational or political ideology identity. In this sense, the scales that measure separate domains are not taking into account the way identities in separate domains may interact and overlap in everyday life. Consequently, these domain-specific scales may be missing an important piece of the puzzle, especially if they hope to generalize to day-to-day life.

Jones and McEwen (2000) suggested a conceptual model of identity that harmonizes a general, day-to-day identity and breaking identity into smaller, more manageable domains. Their framework suggests that individuals have a core, general identity that is comprised of different, but overlapping domain specific identities. In other words, their conceptualization of identity takes into account that there are separate

domains of identity, but that they all overlap and interact throughout daily life to form one's core sense of self. This framework supports the theory that though there are several different domains in which one can measure identity, a more general, core sense of self can also be of interest. Ultimately, domain specific or general identity could be argued as the main focus of research depending on the research question. For the purposes of this study, the more general sense of identity will be the focus.

A more general measure of identity has several appealing qualities to researchers including efficiency and, in some instances, propriety. First, being able to gather information about an individual's general identity would be much less time consuming than gathering information about an individual's identity in several different domains. For instance, in considering large scale testing and the burden placed on the participants, one measure of general identity would be much more efficient than several domain-specific measures of identity. If participants are required to complete a battery of instruments in several different domain areas of identity, their scores could possibly be affected by testing fatigue. Testing fatigue occurs when participants have must complete several instruments and have difficulty maintaining focus and attending to the task at hand. The quality of the responses from participants experiencing testing fatigue begins to decrease as the number of tests increases. In other words, testing fatigue introduces unnecessary measurement error into participants' responses. Ideally, constructs are measured with as little measurement error as possible. To avoid introducing measurement error introduced due to testing fatigue, researchers would be forced to measure a subset of the possible domains of identity. This would lead to subjective decisions by each

researcher as to what domains of identity are the most important to be measured. Instead, a measure that attempts to directly measure general identity would be more efficient.

Second, it may not be appropriate to assume that one's general identity is the sum of its parts (domains). For instance, it is possible that individuals weight some domain identities more heavily than others. It is also possible that several domains overlap in some areas (Jones & McEwen, 2000). For example, it is difficult to imagine that one's religious identity and one's political ideology can be mutually exclusive due to the fact that they are both often based on one's values and beliefs. Whether domains of identity are weighted differently or whether they overlap, it would be inappropriate to assume that adding individuals' separate domain identities would be equivalent to their general sense of self. Thus, a single scale to measure general sense of self, without reference to any specific domain, would be more appropriate.

Sense of Identity Scale

One instrument that addresses one's general sense of self is the Sense of Identity scale developed by Lounsbury and Gibson (2011). It is an 8-item scale on which participants respond to each item using a 5-point Likert rating from "strongly disagree" to "strongly agree." The Sense of Identity scale produces a total score that represents a continuous measure of individuals' general identity. The authors define sense of identity as "knowing one's self and where one is headed in life, having a core set of beliefs and values that guide decisions and actions; and having a sense of purpose." This sense of identity is undeniably similar to Jones and McEwen's (2000) core sense of self. It is also important to note that the Sense of Identity scale seems to align most closely with Marcia's identity achievement stage of identity development. In fact, it has been shown

to correlate most highly with the identity achievement stage (Lounsbury, Huffstetler, Leong, & Gibson, 2005).

The Sense of Identity scale is ideal for large scale research particularly within a higher education setting. An instrument to measure individuals' core sense of identity would be more generalizable to day-to-day student life and thus would be incredibly useful for research in higher education. Administrators may want to use students' general sense of identity to predict performance in several different domains (e.g., academic, behavioral, occupational). Past research has examined how domain specific identity can predict performance in these areas, but as previously discussed, most individuals may not identify with one specific domain identity. Thus, it would be useful to examine how one's general sense of identity can predict performance in these specific domains. The Sense of Identity scale is also ideal for large scale testing due to the fact that it is a very brief scale and thus can easily be given to a large sample of students without concern for testing fatigue. This short, general measure of identity would provide a general snapshot of individuals' identity at a given point in time.

Although the potential benefits of the Sense of Identity scale are clear, validity evidence must be examined before researchers can be confident in the inferences drawn from the scores. It is important to note that, to date, the only validity evidence for the Sense of Identity scale has been collected by the creators of the scale. Specifically, the creators examined the external validity of the scale by investigating whether the scores on the scale related to external variables as would be expected by theory and previous literature. The creators found that the Sense of Identity scale related to several variables as expected, including but not limited to: GPA, intention to withdraw from college

(negatively), satisfaction with social life, satisfaction with safety and security, satisfaction with degree progress, satisfaction with their major, and overall life satisfaction (Lounsbury & Gibson, 2011) .

Correlates/Importance of Identity

If an instrument to measure identity is to be used in research settings, it is necessary to consider the importance of the construct of identity. One way of examining the importance of identity is to examine the relationships identity has with other important variables. Researchers have examined how one's identity relates to attitudinal, academic, and behavioral outcomes. An overwhelming amount of research has been conducted that demonstrates that identity achievement is positively related with numerous desirable attitudinal and academic outcomes. More specifically, a strong sense of identity has been shown to be related to general life satisfaction (Lounsbury, Saudargas, Gibson, & Leong, 2005), collegiate academic achievement (Lounsbury, Huffstetler, Leong, & Gibson 2005), academic motivation (Faye & Sharpe, 2008), career decidedness, optimism (Lounsbury, Saudargas, & Gibson, 2004), self-monitoring, ego-resiliency (Grotevant, 1987), autonomy, reflection, self-esteem, post conventional moral reasoning, mature intimacy, cultural sophistication, and an internal locus of control (Marcia, 1980).

Several studies have also examined the relationship between identity and behavioral outcomes. Toder and Marcia (1973) found that when there was conformity pressure for women, identity achievers were the least likely to conform whereas individuals in the identity diffusion stage were most likely to conform under pressure. Adams et al. (1985) obtained somewhat similar results in that they found that identity

achievers reported conforming for achievement gains whereas identity diffusers reported conforming due to peer pressure. Jones and Hartmann (1988) examined the relationship between identity status and substance use and, interestingly, found that foreclosures reported the lowest frequency of use of cigarettes, inhalants, alcohol, marijuana, and cocaine. They also found that identity diffusers were two times more likely to have tried cigarettes and alcohol, three times as likely to have tried marijuana, four times more likely to have tried inhalants and five times more likely to have tried cocaine than those in the foreclosure group. Lounsbury, Saudargas, and Gibson (2004) examined the relationship between personality traits and students' intention to withdraw from college and found that sense of identity was significantly negatively related to one's likelihood to withdrawal. It is undeniable that individuals with a stronger sense of self tend to be in better attitudinal, academic, and behavioral standing than those who have a less developed identity.

Identity in Higher Education

Because identity has been shown to be related to many positive outcomes, it is not surprising that many higher education institutions have taken interest in the construct. Assessment and accountability movements throughout the past few decades have brought an intense examination of student learning outcomes. Specifically, institutions are required to demonstrate that students are learning the material necessary to meet requirements of a general education program. Universities are held accountable to ensure that every student should graduate with a certain foundation of general education knowledge. Thus a vast majority of the assessment at the university level has been focused on the knowledge-based, cognitive components resulting from a college

education. Some researchers, however, have proposed that these cognitive abilities should not be the only outcomes that are important for students to develop. These researchers argue that higher education institutions should be measuring other, non-cognitive, constructs to show growth in their students. For example, Chickering (1999) suggested that personal qualities and human development should be products of the higher education experience. A vast majority of institutions have programs and organizations that help to foster the growth and development of personal characteristics. Chickering goes on to give examples of some of the most common personal qualities that institutions tend to be interested in, sense of identity being one of the qualities on the list.

Universities nationwide include identity as one of the desired outcomes for undergraduate education. For example, a report developed by the American Association for Higher Education, American College Personnel Association, and the National Association of Student Personnel Administrators entitled *Powerful Partnerships: A shared responsibility for learning* (1998) supported the idea that sense of identity should be a goal for undergraduate education. Baxter Magolda (2003) proposes that a key process of learning should be sharing experiences that shape identity, thus encouraging programs to help foster identity development. Baxter Magolda (2003) highlights the fact that once an individual has encountered and worked through the point in life where the ideals of external authorities clash with internal ideals of the self, they are in a better place to make adult decisions. According to Baxter Magolda, the ability to guide decisions using an internal sense of self instead of relying on external influences such as peer pressure is essential for successful functioning in the real world.

One important objective for many higher education establishments is to prepare students to excel in the work force. As Baxter Magolda (2003) emphasizes, a strong sense of identity is necessary for effectively functioning throughout life, especially after graduation. Thus a strong sense of self would be essential for employment success. Klemp (1977) found that one's knowledge in a specific domain is unrelated to exceptional performance in one's career, but that one's willingness to learn and interpersonal skills are the qualities that distinguish exceptional employees from the rest. The fact that employers weigh personal attributes more heavily when identifying exceptional employees, indicates an undeniable need to measure and develop these characteristics. If universities can help to foster desirable non-cognitive attributes, students may ultimately be more employable after graduate.

It is undeniable why higher education institutions would want to further examine identity development as a desirable outcome of higher education. The attributes universities define as important to foster throughout the college career should be assessed just as the cognitive domains are assessed. In this sense, information from these assessments can be used to help create or improve programs to develop these qualities. In order for institutions to assess human development and personal qualities, three challenges must be met. Administrators must first determine which specific elements should be outcomes of students' experience at their institution. Institutions can then focus on the qualities they feel are most beneficial for students to develop throughout their college career. Programs and/or interventions within an institution that should help to foster growth of personal qualities and human development must then be identified. Instruments must be selected or developed to measure the outcomes outlined by the

administration. As long as the instruments chosen or developed do an adequate job of reflecting the construct, students' scores on these instruments can be used to help inform as to the effectiveness of the programs. Again, knowledge about the effectiveness of programs can help to develop and improve programs that aim to foster development of important outcomes.

Recent emphasis on college student identity formation may lead to changes in policy regarding college students' experiences. For example, it may be beneficial for advisors to suggest that students take the time to explore several content areas early in their college career. Exploration of different content areas would allow students to gain a better idea of content areas they can and cannot identify with, thus assisting them in the formation of their own identity. Baxter Magolda (2003) indicates that it may be beneficial to make the self a central part of learning. She gives four examples of how to promote identity as a central part of learning in multicultural education, community development, academic advising, and teaching. It is unmistakable that with the acceptance of identity as a desirable outcome of higher education, more programs and interventions are likely to be developed to help facilitate the development and formation of students' identity.

Growth/Change Over Time of Identity

If universities hope to nurture identity development it is essential that there is evidence to suggest that the construct can change over time. If identity is a trait-like construct, and thus stable over time, it would be futile to develop programs that focus on attempting to change it. Notably, numerous researchers and, including Erikson, have gathered ample support for the notion that identity should develop and change over time, particularly within the late adolescent years. According to Erikson, most individuals

resolve their industry vs. inferiority crisis and enter into their identity vs. role confusion crisis during adolescence. Specifically, the identity crisis often occurs from puberty throughout the college years (Erikson, 1959). Archer (1982) examined differences in identity formation between sixth, eighth, tenth, and twelfth grade students. She found that, as expected, identity achievers and sophisticated decision making were much more frequent among students later in their adolescent years than among the early adolescents.

Waterman (1982) reviews several studies examining the timing of identity development and notes that the college years seem to be the period in which the largest gains in identity formation occur. Conceptually, this is logical in that attending a university is typically the first occasion in which individuals are not living with their parents (or parental figures). As a result, they are not consistently reinforced based on the beliefs and values of an authority figure, allowing for an opportunity to explore diverse ideals and experiences. In other words, college campuses facilitate unique identity development by exposing students to people, cultures, and life issues that many students have not experienced throughout their early adolescent and high school years.

Longitudinal Research in Identity

Because identity is widely accepted as a developmental process, a fair amount of longitudinal research has been conducted to examine how identity status changes over time. Meeus (2011) provides a thorough review of longitudinal research conducted within the identity literature between 2000 and 2010. The longitudinal studies included in this review add a great deal of information to the domain of identity research, but several of the studies reviewed have limitations that have been previously discussed. First, a vast majority of the studies conducted, use Marcia's paradigm to examine identity status.

Although Marcia's paradigm is well-supported, its limitations from a measurement standpoint (e.g., categorizing individuals) still poses an issue. Second, many of the studies examine identity within specific domains of identity (e.g., occupation, religion, political ideology, etc.). As previously discussed, identity within specific domains can be useful in some conditions, but it may also be important to look at individuals more general, core. Examination of individuals' general sense of identity it will allow researchers to examine how one's day-to-day identity relates to important external variables. Researchers and universities can also use this information to help create and evaluate programs to promote identity development.

CHAPTER IIB: Residuals

As briefly discussed in Chapter I, each technique for analyzing longitudinal data makes assumptions about the residual variances and covariances. Thus, before analyzing any identity data, it is important to thoroughly explore the concept of residuals and how they vary. In order to most effectively demonstrate what residuals are, an example data set with one predictor, t , and a dependent variable, interest, will be used. The values of t range from 1 to 3 and interest can range from 4 to 28. This section first discusses residuals in a traditional regression model, with t treated as a between-subjects continuous predictor. Next, residuals are considered in a regression model with t treated as a between-subjects categorical predictor. Residuals are then discussed in terms of repeated measures data by treating t as a within subjects variable.

Traditional Regression

Consider a situation in which researchers have 705 observations in a data set with predictor, t , and dependent variable, interest. Again, t takes on values from 1 to 3 and interest scores can range from 4 to 28. For clarity, the scores for the first 12 individuals have been provided in Table 1.

Table 1
First 12 Example Individuals for Traditional Regression

t	Interest (y)
1	23
2	15
3	21
1	21
2	14
3	14
1	22
2	23
3	22
1	22
2	19
3	20

In order to examine the relationship between t and interest, researchers decide to estimate a simple regression model, as shown in the equation below. In this equation, y_i is the predicted value of interest for person i , β_0 is the value of interest when t is equal to zero, β_1 is the amount of change in interest for every unit change in t , t_i is the value of t for person i , and e_i is the error (also known as the residual) for person i .

$$y_i = \beta_0 + \beta_1 t_i + e_i \quad (1)$$

After estimating the model, researchers find that the intercept (β_0) is 23.75 and the slope (β_1) is -1.76, indicating that the typical interest score for an individual when t is equal to zero is 23.75 and for every unit increase in t , there is a 1.76 decrease in interest. In addition to the intercept and slope parameters, an error variance (σ^2) of 24.17 is estimated. This error variance indicates the amount of variability in interest scores that cannot be explained by t . The value of the residual for a given individual is simply the observed minus the predicted values, as shown by Table 2 below for the first 12

individuals. The individual residuals are used to compute the overall error variance as shown in Equation 2 below.

$$s^2 = \frac{\sum (X - \bar{X})^2}{n} \quad (2)$$

Table 2
Predicted Scores and Residuals for First 12 Example Individuals

t	Interest	Predicted Interest	Residual
1	23	21.99	1.01
2	15	20.22	-5.22
3	21	18.45	2.55
1	21	21.99	-0.99
2	14	20.22	-6.22
3	14	18.45	-4.45
1	22	21.99	0.01
2	23	20.22	2.78
3	22	18.45	3.55
1	22	21.99	0.01
2	19	20.22	-1.22
3	20	18.45	1.55

Plotting the observed and predicted values on a graph allows the residual term to be visually examined to aid in the explanation of what it represents. Figure 2 below presents individuals' observed scores on the interest variable, as well as their predicted scores based on the regression model. The observed scores are indicated by the diamonds, and the predicted scores are indicated by the line. The graph clearly depicts that the residual value is simply the distance (or difference) from the observed value to the value predicted by the regression model we specified.

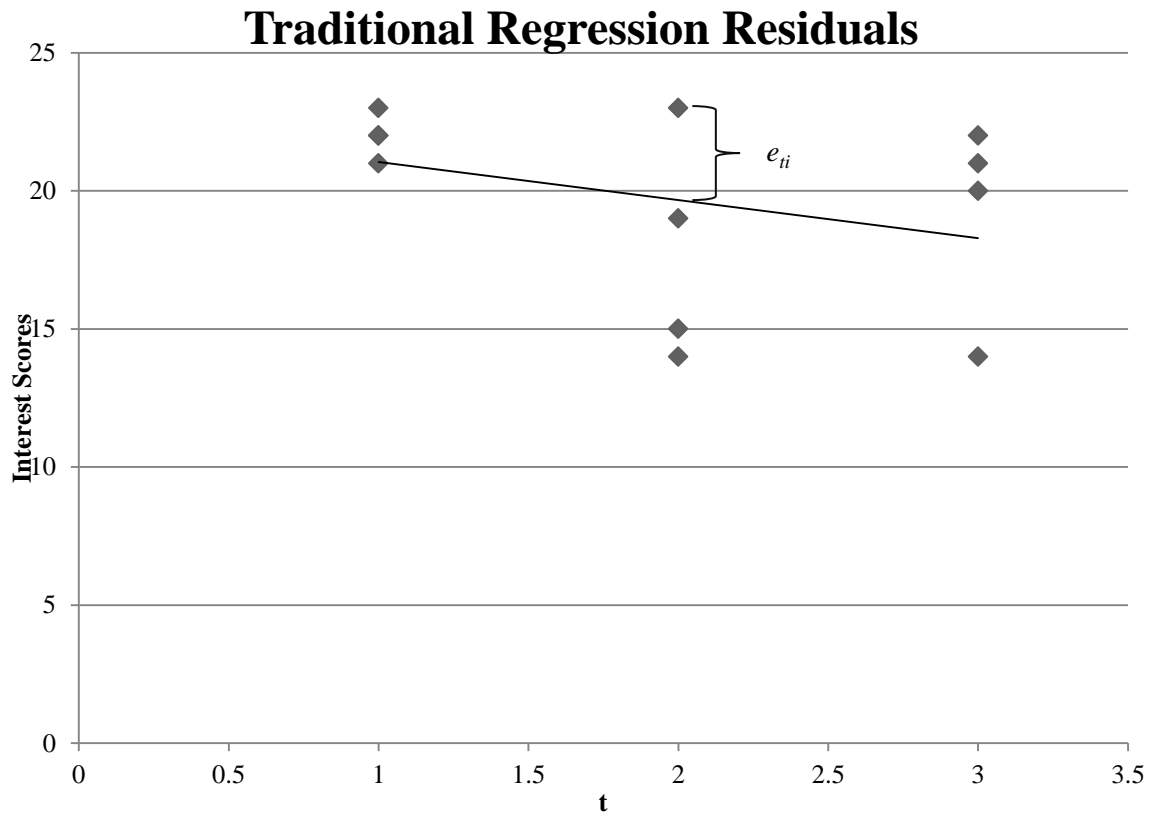


Figure 2. Residual values in traditional regression

Although every statistical technique makes assumptions about the data, the assumption of importance for this explanation is the assumption the traditional regression model makes about the residuals. This assumption states that the residuals in the model are normally distributed with a mean of 0 and a variance equal to σ^2 . This assumption can be written as: $e_i \sim N(0, \sigma^2)$, or in matrix form as $\mathbf{e} \sim N(0, \mathbf{V})$. Thus, “ \mathbf{e} ” represents the vector of errors for all participants, “ N ” indicates that the residuals are normally distributed, “0” indicates that the mean of the errors is zero, and “ \mathbf{V} ” represents the matrix of errors for all participants. If researchers were to write out the \mathbf{e} and \mathbf{V} matrices consisting of

information for all 705 observations, the \mathbf{e} matrix would be 705x1 and the \mathbf{V} matrix would be 705x705. To simplify the presentation, only the first 12 observations of our data are presented for \mathbf{e} and \mathbf{V} in Figure 3 below.

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \\ e_{10} \\ e_{11} \\ e_{12} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 \end{bmatrix} \right)$$

Figure 3. Matrix notation for the assumption of normally distributed residuals

Notably, another regression assumption, the assumption of independent observations, is demonstrated by the \mathbf{V} matrix. This assumption is demonstrated by all zeros on the off diagonal, indicating that the residuals from different individuals are unrelated.

Regression with a Categorical Predictor

Regression can be used not only with continuous predictors, but with categorical predictors. The example data with predictor, t , and dependent variable, interest, will again be used to demonstrate residuals in regression with a categorical predictor. Thus, t in this example will be considered as a nominal, or grouping, variable. Again, t is considered a between subject variable and in this example, the values of 1, 2, and 3 for t indicate three

separate groups. Because t is categorical, two dummy coded variables are often used to represent the variable in the model as shown in Table 3 below.

Table 3

Two Dummy Codes to Represent Categorical Variable t

t	Interest	$t1$	$t2$
1	23	1	0
2	15	0	1
3	21	0	0
1	21	1	0
2	14	0	1
3	14	0	0
1	22	1	0
2	23	0	1
3	22	0	0
1	22	1	0
2	19	0	1
3	20	0	0

Note that participants with a “1” in the $t1$ column are in group 1, participants with a “1” in the $t2$ column are in group 2, and participants with zeros in both columns are in group 3. When two dummy coded variables are used in regression with a categorical predictor variable, the equation can be written as shown below. Note that when $t = 3$, both $t1$ and $t2$ have values of zero; therefore, $t = 3$ is considered the reference group. In this equation, the intercept represents the average interest score for individuals in the group 3, β_1 represents the estimated difference in average interest scores between individuals in group 3 and individuals in group 1, β_2 represents the estimated difference in average interest scores between individuals in group 3 and individuals in group 2, and the residual again represents the difference between the observed score and the predicted score for individual i .

$$y_i = \beta_0 + \beta_1 t_{1i} + \beta_2 t_{2i} + e_i \quad (3)$$

Notably, this model compares the typical interest scores of groups 1 and 2 to the typical interest scores of group 3, but there is no direct comparison of group 1 with group 2. An alternate way to specify a regression model with categorical predictors is to estimate a model with the same number of dummy codes as there are groups (three in this example), but without an intercept. The dummy codes used to estimate a model without an intercept for the interest example are shown in Table 4 below.

Table 4
Three Dummy Codes to Represent Categorical Variable t

t	Interest	t1	t2	t3
1	23	1	0	0
2	15	0	1	0
3	21	0	0	1
1	21	1	0	0
2	14	0	1	0
3	14	0	0	1
1	22	1	0	0
2	23	0	1	0
3	22	0	0	1
1	22	1	0	0
2	19	0	1	0
3	20	0	0	1

The regression equation for three dummy codes without an intercept can be written as shown below. In this equation y_i is individual i 's predicted interest score, β_1 is the average interest score for individuals when $t=1$, β_2 is the average interest score for individuals when $t=2$, β_3 is the average interest score for individuals when $t=3$, and e_i indicates the residual for person i .

$$y_i = \beta_1 t_{1i} + \beta_2 t_{2i} + \beta_3 t_{3i} + e_i \quad (4)$$

This model does not explicitly test differences between group means, but subcommands can be used to examine equality of parameters (and thus equality of group means). When the model without an intercept is estimated, you find that β_1 takes a value of 22.28, β_2 is estimated to be 19.63, and β_3 takes on a value of 18.74, the means of the groups, respectively. Additionally, the estimated error variance (σ^2) is estimated to be 24.02. The square root of this value can be calculated to demonstrate the typical distance of individuals' observed scores from the predicted scores (in this model, the means). Thus, the typical residual value is approximately 4.9. Notably, the models in Equation 3 and Equation 4 are equivalent models, meaning that they produce the same predicted values and the same errors as shown in the table below. Consequently, equivalent models also produce the same model-data fit.

Table 5
Predicted Scores and Residuals for First 12 Example Individuals with Categorical Predictor

t	Interest	Predicted Interest	Residual
1	23	22.28	0.72
2	15	19.63	-4.63
3	21	18.74	2.26
1	21	22.28	-1.28
2	14	19.63	-5.63
3	14	18.74	-4.74
1	22	22.28	-0.28
2	23	19.63	3.37
3	22	18.74	3.26
1	22	22.28	-0.28
2	19	19.63	-0.63
3	20	18.74	1.26

Again the observed and predicted values can be plotted graphically in order to visually examine residual values. Predicted values for individuals in this model are equal to their

respective group means. In the graph below, diamonds indicate observed scores, squares indicate predicted scores. For a given individual, the residual value is the distance between the observed score and the predicted score (which is the mean for their respective group)

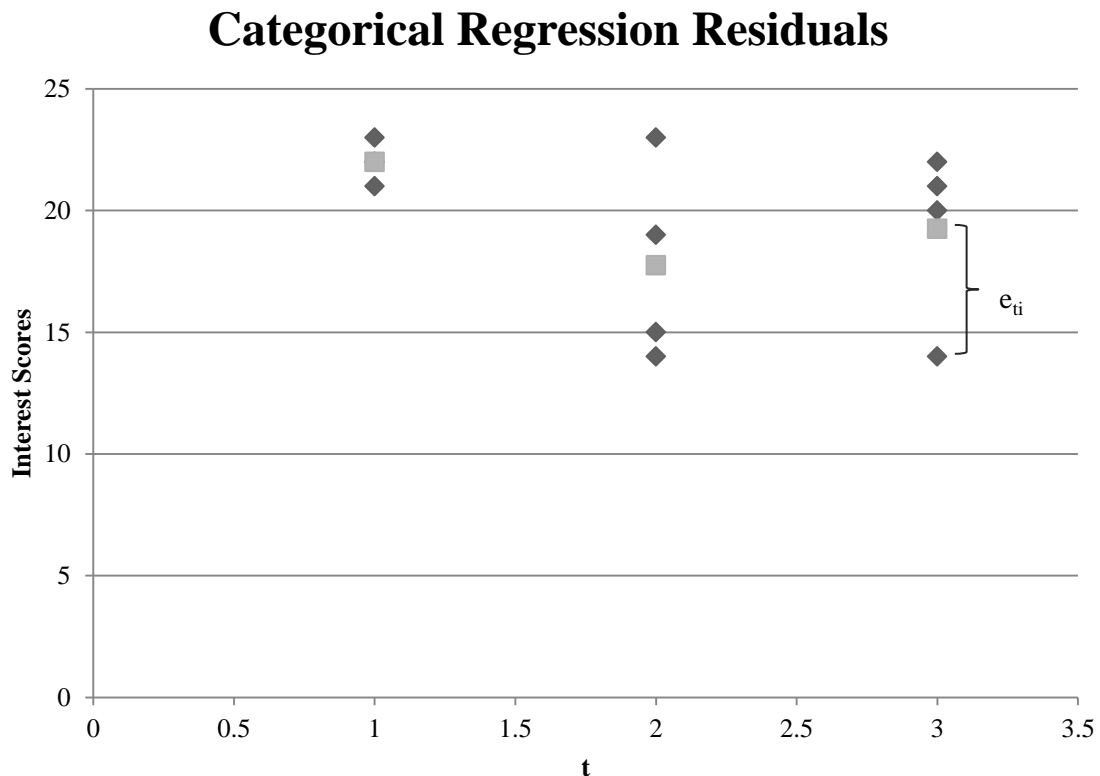


Figure 4. Residual values in regression with a categorical predictor

Two important characteristics of the categorical regression model should be noted, regardless of whether it is estimated with two dummy codes or with three dummy codes and no intercept. First, these models are the same as a one-way between-subjects ANOVA model. Thus, the hypothesis that all means are equal can be tested by comparing the categorical regression model presented above with an intercept only model.

Comparing these two models test the null hypothesis that $\mu_1 = \mu_2 = \mu_3$. The second

characteristic has to do with the error variance. The extent to which residuals vary in a group is equal to the variances of the observed scores in each group. The regression model, however, assumes that residual variance is the same across all levels of predictor. Thus, a single error variance is estimated and is equal to the pooled variance across groups, which is the within group variance in ANOVA. Because in our example there are an equal number of individuals in each of the three groups, the pooled within group variance is simply the average of the three residual variances which equals the observed variance for each group. Note that the variance for group 1 is 15.86, the variance for group 2 is 26 and the variance for group 3 is 30.22. If these three values are averaged, a value of 24.03 is obtained. Note that this value is very similar to the variance estimate when t was treated as a continuous variable. The similar variance components suggest that the model that does not impose a linear model and the model that imposes a linear model have similar predictive ability. It is likely that the linear model produces predicted scores similar to the means at each measurement occasion.

As with traditional regression with continuous predictors, the categorical regression model, and thus the one-way between-subjects ANOVA, makes the assumption that residuals are normally distributed with a mean of zero and a variance of σ^2 . Thus, the matrix form for this assumption is the same as that provided for the traditional regression model (See Figure 3. Notably, the error variance for each observation is σ^2 , indicating that error variances (in this case the pooled within group variance) are equal across groups. This is also known as the homogeneity of variance assumption. Although statistical tests can inform researchers as to the extent to which this assumption

is violated, simple inspection of the variances presented above suggest that variances are not equal across groups and thus a single error variance may not be appropriate.

Repeated Measures Data

With the knowledge of residuals and assumptions about residuals above, it would be useful to revisit repeated measures data. The ongoing interest example is actually repeated measures data in which 235 individuals' interest levels were measured at 3 time points. Thus, the three values of t correspond to the first, second, and third measurement occasions. If the first regression model in Equation 1, where t is treated as continuous, were estimated for this data, it would specify a linear relationship between time and interest scores. The second model in Equations 2 (or 3), where t is treated as categorical, would differ from the first in that it would not specify the form of the relationship between time and interest but would instead model predicted scores at each specific level of the independent variable. However, with repeated measures data, both of these models would be inappropriate due to their violation of the assumption of independent observations. This can be shown by examining the off diagonal of the \mathbf{V} matrix for the first four persons in the repeated measures data below.

$$\begin{array}{l}
 \text{Person 1} \\
 \text{Person 2} \\
 \text{Person 3} \\
 \text{Person 4}
 \end{array}
 \left[\begin{array}{c}
 e_{11} \\
 e_{21} \\
 e_{31} \\
 e_{12} \\
 e_{22} \\
 e_{32} \\
 e_{13} \\
 e_{23} \\
 e_{33} \\
 e_{14} \\
 e_{24} \\
 e_{34}
 \end{array} \right] \sim N \left(\begin{array}{c}
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}, \begin{array}{cccccccccccc}
 \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2
 \end{array} \right)$$

Figure 5. Demonstration of the assumption of independent observations

Note that each residual term has two subscripts, the first for time and the second for person. Thus e_{21} corresponds to the residual for person 1 at time 2. The square with the solid line indicates the residual covariance matrix for one individual, which can also be represented as

$$V_i = \begin{bmatrix} \sigma^2 & & \\ 0 & \sigma^2 & \\ 0 & 0 & \sigma^2 \end{bmatrix} \quad (5)$$

The zeros on the off diagonal indicate that the residuals for each individual are uncorrelated. In other words, an individual's interest score at time 1 is completely unrelated to their interest score at time 2. It is unrealistic to assume that responses coming from the same individual would be completely unrelated at different time points and thus the two regression models (Equations 1, 3, or 4) proposed above that make this assumption would be inappropriate. The square with the dotted line indicates that errors

from person 1 are unrelated to errors of person 2. This demonstrates the independence of observation assumption previously discussed.

The within-subjects regression model can be illustrated using the same data as in Table 4, with the exception of t now being considered a within subjects variable. In order to estimate a within-subjects regression for the interest data treating t as a categorical variable, the model presented below can be specified.

$$y_{it} = \beta_1 t1_{it} + \beta_2 t2_{it} + \beta_3 t3_{it} + u_{0i} + e_{it} \quad (6)$$

In this equation, y_{it} indicates the predicted interest score for person i at time t , β_1 indicates the typical score at time 1, β_2 is the typical score at time 2, β_3 is the typical score at time 3, and e_{it} is the residual for person i at time t . Because t is represented in Equation 6 using three dummy-coded variables, no form is being specified for the relationship between time (t) and interest. As well, the regression coefficients β_1 , β_2 and β_3 will again equal the average interest score for each value of t (e.g., the average interest score at each time point), making this model equivalent to a within-subjects ANOVA.

The additional parameter, u_{0i} , is the “person effect” and indicates to what extent a person’s average deviates from the overall average. For clarity, Table 6 has been provided to demonstrate what the “person effect” is. Note that u_{0i} is simply an individual’s average interest score across the three time points subtracted from the grand mean (the mean across all individuals and all measurement occasions), which is 20.22 in this example. Thus, on average, the first participant’s interest scores are about 4 points lower than the grand mean.

Table 6
Demonstration of the Person Effect

<i>i</i>	<i>y</i>			Person Average	<i>u</i> _{0<i>i</i>}
	t=1	t=2	t=3		
1	21	14	14	16.33	-3.88
2	22	23	22	22.33	2.12

Again, scores can be plotted graphically to help foster understanding of the residual components (e_{ii} and u_{0i}). The graph below presents interest data for the two participants. The diamonds represent the individuals' observed scores, the squares represent the individuals' predicted scores, and the small circles represent the individuals' predicted scores plus the person effects. In other words, the small circles can be thought of as each individual's predicted score when taking into account the "person" effect. Note that in a between subjects model, the residual would simply be the distance from the observed score to the predicted score, whereas in the current model the residual variance is broken down into two parts: u_{0i} and e_{ii} . The distance between the predicted value and the predicted value plus the person effect is u_{0i} and represents the spread of individual predicted scores around the overall predicted score. The distance from the predicted value plus the person effect and the observed value is e_{ii} and represents the spread of individuals' observed scores around their respective predicted scores. Because the model contains two residual terms, two residual variances are estimated: τ_{00} and σ^2 . τ_{00} is the variance for the between-person random effect, u_{0i} , and σ^2 is the variance of the within-person random effect e_{ii} . Conceptually, τ_{00} indicates the extent to which individuals' predicted scores plus person effects (small circles) vary about the overall predicted score for individuals (squares). In essence, this provides information as to how individuals

differ in their interest scores averaged across time points. σ^2 , on the other hand, captures the extent to which an individual's observed interest scores (diamonds) vary about their predicted score when the person effect is taken into account (small circles). Notably, σ^2 is assumed to be the same at each time point.

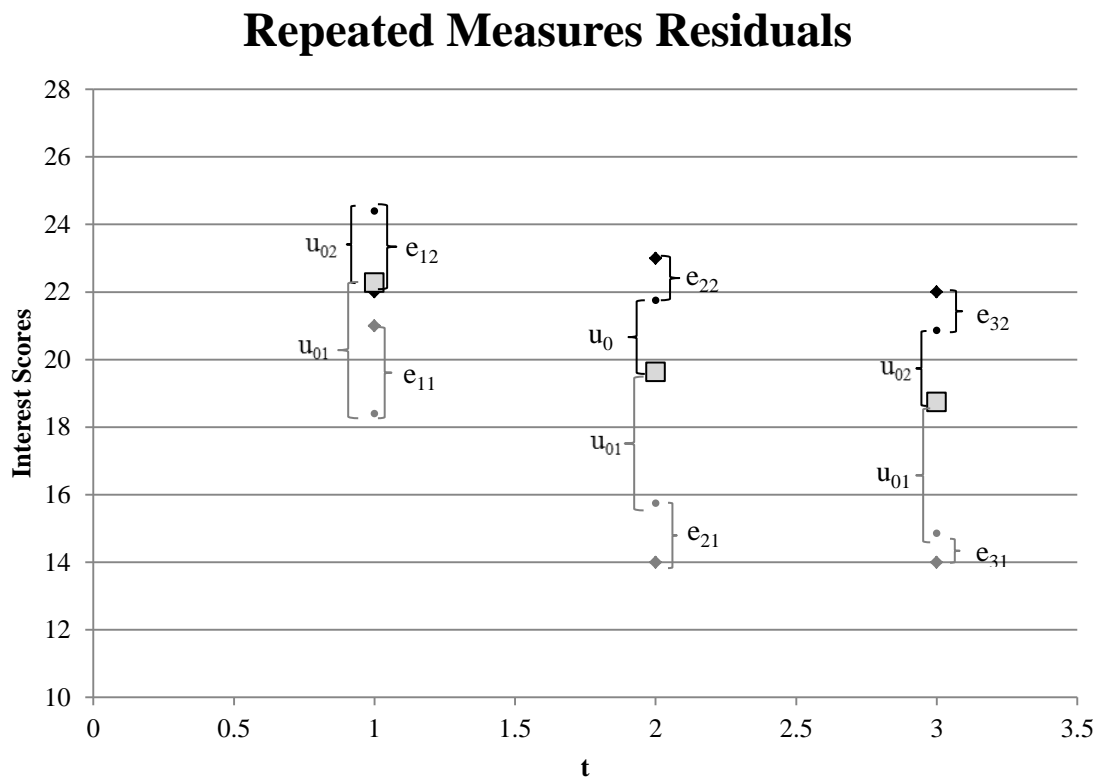


Figure 6. Residual values in repeated measures models

The commonly used within subjects univariate ANOVA assumes that the total residual variation ($\sigma^2 + \tau_{00}$) is the same for each measurement occasion, and that the relationship between all measurement occasions, as indicated by estimation of one τ_{00} parameter. This assumption is the assumption of compound symmetry and is often presented in matrix form as shown below.

$$V_i = \begin{bmatrix} \sigma^2 + \tau_{00} & & \\ \tau_{00} & \sigma^2 + \tau_{00} & \\ \tau_{00} & \tau_{00} & \sigma^2 + \tau_{00} \end{bmatrix} \quad (7)$$

It may seem unusual that τ_{00} is both a variance and a covariance, but in this model τ_{00} captures not only differences among persons in interest scores averaged across time, but also the extent to which interest scores covary within persons.

Because τ_{00} is a covariance, it may be difficult to interpret. For this reason, it is often converted into a correlation to better understand the relationship between measurement occasions within persons. Note that with this example, there are three possible correlations between measurement occasions. This model assumes that a single correlation is sufficient to adequately model the relationship between time 1 and time 2, time 2 and time 3, and time 1 and time 3. In other words, this assumption states that the relationship between measurement occasions is the same for all individuals.

Just as interest variances can be examined in a between subjects ANOVA to determine the plausibility of satisfying the homogeneity of variance assumption, the interest variances and covariances/correlations can be examined in a within subjects ANOVA to determine the plausibility of satisfying compound symmetry. Note that these are the statistics associated with the interest scores at each time point, not the statistics associated with the residuals of the model. The pattern of statistics in the observed covariance matrix can be consulted, however, to ascertain whether a compound symmetric form is appropriate for the residual covariance matrix. The interest variances, covariances, and correlations are displayed in the matrix below. Note that the variances

are on the diagonal, the covariances are on the bottom off diagonal, and the correlations are on the top off diagonal.

$$\begin{bmatrix} 15.86 & .67 & .59 \\ 13.69 & 26.00 & .77 \\ 12.92 & 21.55 & 30.22 \end{bmatrix} \quad (8)$$

Examination of the matrix above will help to provide insight as to whether or not compound symmetry is a plausible assumption for this data. The variances can first be examined to determine whether or not they seem to be constant across measurement occasions. In our example, it seems that variances are not constant across measurement occasions (they range from 15.86 to 30.22). More specifically, it seems that variances increase across time points. The covariances and correlations can be examined to determine whether one value could adequately represent the three values observed in the data. In our example, the covariances and correlations do not seem to be equivalent (the correlations range from .59 to .77). Notably, the correlations between adjacent time points (time 1 and 2, and time 2 and 3) have a higher magnitude than non-adjacent time points (time 1 and time 3). Clearly the observed variances, covariances and correlations suggest that the assumption of compound symmetry for the residual covariance matrix may not be plausible.

As discussed in Chapter 1, the assumption of compound symmetry is incredibly restrictive and, in practice, the more relaxed assumption of sphericity is sufficient when employing repeated measures ANOVA. Thus, ANOVA allows for a “Type H” residual covariance matrix as shown in Table 7. (Henceforth, all residual covariance matrices will be shown in Table 7 and discussed more generally rather than using the interest data.)

Table 7

Residual Covariance Structures

Covariance Structure	Matrix Form	Parameters	Number of Parameters Estimated
Compound Symmetry	$\begin{bmatrix} \tau_{00} + \sigma^2 & \tau_{00} & \tau_{00} \\ \tau_{00} & \tau_{00} + \sigma^2 & \tau_{00} \\ \tau_{00} & \tau_{00} & \tau_{00} + \sigma^2 \end{bmatrix}$	τ_{00}, σ^2	2
Huynh-Feldt (a.k.a. Type H)	$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda & \frac{\sigma_1^2 + \sigma_3^2}{2} - \lambda \\ \frac{\sigma_2^2 + \sigma_1^2}{2} - \lambda & \sigma_2^2 & \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda \\ \frac{\sigma_3^2 + \sigma_1^2}{2} - \lambda & \frac{\sigma_3^2 + \sigma_2^2}{2} - \lambda & \sigma_3^2 \end{bmatrix}$	$\sigma_1^2, \sigma_2^2, \sigma_3^2, \lambda$	4
Toeplitz	$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}$	$\sigma^2, \sigma_1, \sigma_2$	3
Homogeneous Autoregressive	$\begin{bmatrix} \sigma^2 & \sigma\rho & \sigma\rho^2 \\ \sigma\rho & \sigma^2 & \sigma\rho \\ \sigma\rho^2 & \sigma\rho & \sigma^2 \end{bmatrix}$	σ^2, ρ	2
Heterogeneous Autoregressive	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 \end{bmatrix}$	$\sigma_1^2, \sigma_2^2, \sigma_3^2, \rho$	4
Unstructured	$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$	$\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_{12}, \sigma_{13}, \sigma_{23}$	6

The Type H residual covariance matrix corresponds to the assumption of sphericity (Wolfinger & Chang, 1995) and the within-subjects ANOVA most commonly used in SPSS or SAS assumes sphericity. Again, if the assumptions about the matrix employed do not hold, conclusions regarding mean differences can be affected. For this reason, information about the extent to which the observed covariance matrix departs from sphericity is provided in the output (e.g., epsilon, Mauchly's test of sphericity), as

described in Chapter I. If the assumption of sphericity is violated, the researcher has the option of using the results of the within-subjects ANOVA where the degrees of freedom have been adjusted by the degree to which sphericity has been violated. They also have the option of using a MANOVA to examine mean differences in a variable over time.

As reviewed in Chapter I, a MANOVA model assumes an unstructured residual covariance matrix, as shown in Table 7. This matrix requires every parameter to be estimated and thus provides the optimum amount of information about the data. In fact, the unrestricted covariance matrix will equal the observed covariance matrix. Because every parameter is estimated, however, MANOVA may have issues acquiring precise estimates for parameters as well as issues with capitalization on idiosyncrasies in the data. In addition, the degrees of freedom used for the denominator are based on the number of persons, not the number of total observations (each individual has multiple observations), making the denominator degrees of freedom are smaller and possibly increasing the risk of Type II errors (Hoffman, in preparation).

ACS Modeling

The compound symmetric residual covariance matrix and the unstructured residual covariance matrix form the extreme ends of the residual covariance matrix continuum. The compound symmetric matrix, which is very parsimonious and very restrictive, is at one end. At the opposite end is the unstructured matrix, which is much more flexible but also much less parsimonious. It would be beneficial to employ methods that allow residual covariance matrices somewhere in the middle of the continuum. As briefly discussed in Chapter I, ACS modeling allows researchers to specify many

different residual covariance matrices, including but not limited to the compound symmetric and unstructured residual covariance matrices.

Like within-subjects ANOVA or MANOVA, ACS model can be used when the interest is in comparing the means of a variable over time. ACS models can also be used to model other forms of the relationship between time and the dependent variable, such as a linear or quadratic relationship¹. ACS modeling has the benefit over within-subjects ANOVA or MANOVA because it not only allows for specification of several residual covariance matrices, but also allows for analysis of Type II data. The allowance of Type II data and specification of several residual covariance matrices offers immense advantages over traditional repeated measures ANOVA and MANOVA.

Chapter I presented one example of a “happy medium” covariance structure that can be used with ACS modeling known as the homogeneous autoregressive covariance matrix. Recall that this matrix assumes that all residual variances are equal and that adjacent time points are more strongly related than non-adjacent time points. As shown in Table 7, two parameters need to be estimated for this residual covariance matrix: σ^2 and ρ . Note that the adjacent covariances shown in Table 7 are calculated by multiplying the variance by the relationship between measurement occasions (ρ). For the non-adjacent time points that are one step away from each other, the variance is multiplied by ρ^2 . If time points were three steps away from each other (e.g., time 1 and time 4), the variance would be multiplied by ρ^3 , and so on. This residual covariance structure would be most appropriate when researchers feel that the variability of the scores over time is stable, but that adjacent time points are more related than non-adjacent time points.

¹ Although not reviewed in this chapter, linear, quadratic or cubic trends can also be investigated in the context of within-subjects ANOVA. It is more common, however, for within-subjects ANOVA to only be used to assess

There is also an option to have heterogeneous residual variances with the autoregressive covariance matrix, as shown in Table 7. Thus, with three time points, four parameters would need to be estimated: all three residual variances and ρ . This is still more economical than six parameters being estimated as in an unstructured covariance matrix, but not as parsimonious as only two parameters being estimated as in the compound symmetry or homogeneous autoregressive covariance matrices. Again, this residual covariance structure would be useful in situations in which researchers believe the residual variances differed and residual covariances for adjacent time points are larger than those for non-adjacent time points.

Another example of a “happy medium” residual covariance structure is the toeplitz residual covariance matrix (see Table 7). Similar to many of the other matrices, it assumes that the variances of the residuals are equal across time points. It is similar to the autoregressive structures in that the toeplitz residual covariance matrix assumes residual covariances between adjacent time points are equal, but differs slightly in that it does not restrict the residual covariances of adjacent time points to be *more* related than nonadjacent time points. With three time points, this would mean that the residual covariances between time points 1 and 2 and time points 2 and 3 would be forced to be equal and the residual covariance between time points 1 and 3 would be different, as in autoregressive. Unlike autoregressive, however, the residual covariance between time points 1 and 3 is not constrained to be systematically smaller. Thus with three measurement occasions, three parameters must be estimated: the residual variance, the residual covariance for adjacent measurement occasions, and the residual covariance for the non-adjacent measurement occasions. This residual covariance structure would be

most appropriate in situations where researchers believe that the residuals follow a similar pattern as the autoregressive residual covariance matrix, but do not feel that nonadjacent measurement occasions need to be constrained to be systematically less related than adjacent measurement occasions. The toeplitz residual covariance structure is more economical than the unstructured residual covariance matrix, but not quite as parsimonious as the compound symmetry or homogeneous autoregressive residual covariance matrices.

Choosing Among Covariance Matrices

As previously discussed, ACS Modeling allows researchers to fit a variety of different models, each with a different residual covariance structure, to the data. Fit indices can then be used to decide which model and residual covariance structure best reflects the data. Thus, it is important to consider how models with different residual covariance structures are chosen or rejected. Ultimately, researchers should base their decision first and foremost on theory and what theory would suggest about how the residual variances and covariances should behave. The examination of the observed covariance or correlation matrix can then be used as supplemental evidence for the theoretical decision as to which residual covariance structure should be used. In a more exploratory situation, examining the observed relationships can also aid in the identification of plausible residual covariance structures, particularly when theory is in its initial stage. It is also possible that theory cannot differentiate between some of the residual covariance structures. For example, homogeneous autoregressive and toeplitz are very similar. In this case it would be helpful to perform separate analyses for both residual covariance structures to empirically examine which structure may be more

appropriate. Again, it is imperative that the researcher first identify plausible residual covariance matrices based on what theory would expect and supplement this information empirically. The selection of an appropriate residual covariance matrix is crucial in obtaining dependable results.

In order to empirically compare models, fit indices are utilized. Unlike the within-subjects ANOVA or MANOVA commonly used in SPSS or SAS, ACS models are estimated using maximum likelihood estimation. Maximum likelihood estimation provides three fit indices for ACS models. The $-2 \log$ likelihood, or deviance statistic, can be used as an indication of model fit. The deviance values in and of themselves are of little interest; however, the comparison of deviances across models is informative. When models are nested within one another, likelihood ratio tests comparing the deviance statistics can be employed to determine if the more parsimonious model fits significantly worse than the more complex model. By definition, more complex models will always fit the data the best, however if a more restrictive model fits just as well or is not significantly worse, it is often considered the more desirable model. Regardless of whether models are nested, the two other fit indices, the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC), can be used to help compare models. The AIC and BIC penalize models for complexity, and thus the values can be compared across nested or non-nested models to determine the most appropriate residual covariance matrix.

Limitations of ACS Modeling

As previously mentioned, ACS modeling offers advantages over traditional repeated measures ANOVA and MANOVA because it allows Type II data and flexibility

in specifying residual covariance structures. The focus of ACS modeling, however, is the similar to repeated measures ANOVA and MANOVA in that it is concentrated on the overall relationship of time with the dependent variable and not individual differences in this relationship.

As an example, consider testing a linear trend in repeated measures ANOVA or specifying a linear relationship between time and the dependent variable in ACS modeling. The limitation for these models is that they only provide information about overall change. In Chapter 1 an example was used in which individuals changed on a construct differently (some increased, some decreased), but collapsing across individuals it seemed as if there was no overall change. Thus, if repeated measures ANOVA, MANOVA, or ACS modeling was employed in this situation, the results would suggest that there is no change on the construct over time while, in reality, individuals are changing quite a bit, but with different trajectories. Notably, specification of the compound symmetric residual covariance matrix is the only model that explicitly states that there are no slope differences across individuals and only examines intercept differences. This is due to the fact that the compound symmetric residual covariance structure posits that the residual covariances between measurement occasions are equivalent. This assumption suggests that individuals have the same trajectory and thus do not change in rank order over time. Notably compound symmetry posits that the only reason interest scores are related across measurement occasions are due to “constant mean differences over time” (Hoffman, in preparation, p. 9). The other residual covariance matrices do allow for differences in slopes, however, they do not provide parameters that easily allow for a discussion of individual differences in change over

time. One method that *can* focus on individual change over time is multilevel modeling, as discussed in the following sections. The first section describes multilevel modeling in general and the second section describes multilevel modeling in the context of longitudinal data.

Multilevel modeling (MLM), also known as hierarchical linear modeling (HLM), is a regression technique used with data in which the assumption of independence of observations is violated. The classic example used to describe this violation of independent observations is students nested within schools. As an example, consider a researcher collecting math scores from students from ten different schools. It would be expected that scores from two students in the same school would be more alike than the scores from two students in different schools. Students from the same school would be more alike because they are in the same environment, have the same teachers, and interact with each other, whereas two students in different schools do not have these similarities. Thus, it would be inappropriate to assume all observations are independent because some are clearly more related than others. If the school effect, or the dependency due to observations being nested in to the same school, is not taken into account the assumption of independent observations will be violated and the standard errors will be underestimated. Underestimated standard errors can lead to increased Type I errors. MLM takes into account the school effect due to students (level one) being nested within schools (level two) using a 2-level model and thus is appropriate when data is nested.

Multilevel Models for Longitudinal Data

The same issue with independence of observations appears in longitudinal data analysis. In longitudinal analyses, an individual is measured at several time points. Thus,

the data is nested in that measurement occasions are nested within people. As previously discussed, it would be inappropriate to assume that one observation from one individual would be independent from another observation from the same person. Just as MLM is able to take into account the school effect in the example above, with longitudinal data MLM is able to take into account the person effect due to measurement occasions (level one) being nested within persons (level two) using a 2-level model.

When employing MLM, a series of models are fit to the data, which typically begins with two models: the unconditional means model and the unconditional growth model. Each of these models contains two levels of information. The first level contains information about individuals including the estimated parameters for the individual as well as within-person variation. Within-person variation refers to the variability of individuals' scores around their own predicted trajectory. Recall that within-person variation is captured by the σ^2 parameter. The second level contains overall information about persons in the population as well as information regarding between-person variability. Between-person variability refers to the variability of individual's predicted scores around the overall predicted scores. Again, recall that between-person variation is captured by the τ_{00} parameter. In order to most clearly describe the unconditional means and unconditional growth models, an example in which sense of identity is measured at three time points will be used.

Before introducing the models, a distinction between the residual covariance matrix used with the previously introduced technique and that used with MLM is needed. With previous techniques we have only examined one residual covariance matrix, which contains all residual variability and covariability. We called this matrix the \mathbf{V} matrix.

MLM also has a \mathbf{V} matrix and it also contains all residual variability and covariability. However, in MLM the \mathbf{V} matrix is a combination of the level one and level two residual covariance matrices. The level one residual covariance matrix in MLM is the \mathbf{R} matrix, and contains information about within-subject variability. The level two residual covariance matrix in MLM is the \mathbf{G} matrix and provides information about the between-subjects residual variance. In traditional techniques, no distinction was made between within and between subjects variability and all residual variability was contained in a single matrix; essentially, $\mathbf{V} = \mathbf{R}$ with traditional methods. Thus in MLM, there is also a \mathbf{V} matrix, but it consists of information from both \mathbf{G} and \mathbf{R} . The combination of \mathbf{G} and \mathbf{R} to form \mathbf{V} will be demonstrated in the following sections.

Unconditional means model. Level one of the unconditional means model captures each individual's mean level of sense of identity across time and within-person variability in sense of identity scores across time from this person average, whereas level two captures the overall mean of the sense of identity scores across people and between-person variability in sense of identity from this group average. The two levels can be represented by the following two equations:

$$y_{it} = \pi_{0i} + e_{it} \quad (9)$$

$$\pi_{0i} = \beta_{00} + u_{0i} \quad (10)$$

The first equation represents the level one or the time level of the model, where y_{it} is individual i 's sense of identity score at time t , π_{0i} is individual i 's intercept and e_{it} is the residual, or how much individual i 's score deviates from their intercept at time t . The second equation represents level two or the person level of the model, where π_{0i} is still

individual i 's intercept, β_{00} is the overall intercept across people, and u_{0i} is the residual, or how much each individual's intercept deviates from the overall intercept. Notably when no predictors are in the model the individual intercept is simply the individual mean (i.e., an individual's average sense of identity score across measurement occasions) and the overall intercept is the overall mean (i.e., the average sense of identity score across occasions and persons).

Within-person variation in sense of identity scores across time is captured by the variance of e_{ti} , denoted σ^2 . A large σ^2 estimate indicates that there is a sizeable amount of within-person variability and thus adding time-varying predictors², such as time, could help to explain within-person variation in scores.

The second level in the unconditional means model estimates an overall mean intercept as well as the variability of the intercepts. A large variance estimate for the level two residuals, denoted as τ_{00} , indicates that individuals have considerably different means for their sense of identity scores. In other words, there is a lot of unexplained variability in mean scores on sense of identity. Adding time-invariant predictors³, such as whether or not individuals participated in a program to develop identity, may help to explain this variance.

Notably, the two equations presented as Equation 9 and Equation 10 can be rewritten as one equation as shown below:

$$y_{ii} = \beta_{00} + u_{0i} + e_{ti} \quad (11)$$

In this equation there are two residual components, u_{0i} and e_{ti} , indicating that the residual variance has been partitioned into within and between-person variation. Note that this

² A time-varying predictor is a variable whose values change across time.

³ A time-invariant predictor is a variable whose values remain stable across time.

equation is very similar to the equation used with repeated measures regression/repeated measures ANOVA. This equation only estimates the grand mean while repeated measures ANOVA estimates the means for all three groups, but the variance is partitioned into the same two parts. Consider a situation where τ_{00} equals 0. In this scenario, all variability in sense of identity scores is within persons, not between persons. In other words, there is no effect for persons and u_{0i} can be dropped from the model, making it a traditional between-subjects regression model. Thus, the dependency imposed by observations being taken from the same person is taken into account by including u_{0i} , which is the person effect in the model.

Unconditional growth model. The unconditional growth model is similar to the unconditional means model in that the first level contains parameters for the individual and captures within-person variability, whereas the second level contains overall parameters and captures between-person variability. The difference between the unconditional means model and the unconditional growth model is the inclusion of time as a level one, or time-varying, predictor. The inclusion of time allows for a slope parameter to be estimated. Because two parameters, a slope and an intercept, are included in the level one model, two equations are used in the level two model. The level one and two equations are as follows:

$$y_{it} = \pi_{0i} + \pi_{1i}t_{it} + e_{it} \quad (12)$$

$$\begin{aligned} \pi_{0i} &= \beta_{00} + u_{0i} \\ \pi_{1i} &= \beta_{10} \end{aligned} \quad (13)$$

The first equation (shown as Equation 12) is the level one model and is similar to the level one model in the unconditional means model in that y_{it} is the predicted score for individual i at time t , π_{0i} is the individual's intercept, and e_{it} is the residual, or the deviation of each individual's observed score from their predicted score. In this model, however, the intercept takes on a different interpretation and represents an individual's sense of identity when time is equal to zero. Notably, the coding of time can change the interpretation of the intercept value. For example, if the initial time point is set to 0, then the intercept would equal an individual's sense of identity score at the initial time point. The new term estimated in this equation is the π_{1i} term or the slope for each individual. The slope indicates the amount of change in sense of identity for each unit change in time. Again, the coding of time (e.g., days, months, years) can alter the interpretation of the slope parameter. The residual, again, indicates how much of the within-person variation is left unexplained by time. If there is a sizeable amount of unexplained variance, other models that include additional time-varying predictors may help to explain the remaining variation. Notably, if time is a strong predictor of sense of identity scores, σ^2 for the unconditional growth model will be smaller than the unconditional means model. The difference in residual variation between the two models can be examined using a pseudo R^2 statistic which indicates the proportion of level one variation explained by time.

Level two of the unconditional growth model is, again, similar to the unconditional means model in that the first equation is exactly the same as the level two equation for the unconditional means model. The overall mean intercept, however, takes on a different interpretation and represents, on average, participants' sense of identity

scores when time equals zero. The residual for this equation indicates how much an individual's intercept deviates from the overall intercept and τ_{00} represents between-person variation in intercepts. Because a slope parameter is estimated with the unconditional growth model, the second equation estimates an overall slope, which indicates, on average, how much participants' sense of identity scores change per unit change in time. Notably the residual term for the slope parameter for this model is constrained to be zero. This postulates that every individual has the exact same slope parameter. Although a residual value can be specified which would allow slopes to vary, it has been omitted from this model to demonstrate how the **G** and **R** matrices can combine to form a familiar residual covariance structure. The **R** matrix in MLM is most commonly assumed to have a homogenous independence form, whereas the **G** matrix has an unstructured form. In this example, the **G** matrix consists of only τ_{00} because only intercepts are permitted to vary. Thus, the **R** and **G** matrices combine as shown below:

G	R	V= ZGZ'+R
$[\tau_{00}]$	$\begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$	$\begin{bmatrix} \tau_{00} + \sigma^2 & \tau_{00} & \tau_{00} \\ \tau_{00} & \tau_{00} + \sigma^2 & \tau_{00} \\ \tau_{00} & \tau_{00} & \tau_{00} + \sigma^2 \end{bmatrix}$

Figure 7. Formation of the **V** matrix from the **G** and **R** matrices⁴

As shown, the unconditional growth model with random intercepts and fixed slopes results in a compound symmetric residual covariance matrix. This model is the same as repeated measures ANOVA assuming a compound symmetric residual covariance matrix, but with time treated as a continuous variable as opposed to a categorical variable (e.g., testing for a linear trend). The similarities are even more

⁴ **Z** is a design matrix indicating which effects are random. In this model **Z** is a 3x1 vector consisting only of ones since only slopes randomly vary.

evident by comparing the repeated measures ANOVA model in Equation 4 to Equations 11 and 12 written as a single equation.

$$y_{ii} = \beta_{00} + \beta_{01}t_{ii} + u_{0i} + e_{ii} \quad (14)$$

The model is also equivalent to an ACS model with time as a continuous predictor and a compound symmetric residual covariance structure. Although all three models are equivalent, note how the focus in a linear trend analysis in repeated measures ANOVA and the ACS model is on the significance of the linear relationship between time and dependent variable. This relationship is the overall relationship. Although individual differences in intercepts are specified in all three models, τ_{00} is a parameter that is only interpreted in the MLM approach. In the ANOVA and ACS modeling approach, τ_{00} might not even be reported, much less interpreted.

The fact that all these models are equivalent highlights a restrictive assumption made by repeated measures ANOVA with a linear trend and an ACS model with a compound symmetric residual covariance structure, which is that individual trajectories all have the same slope and only vary in their intercepts. However, in MLM researchers do not have to constrain the slopes to be equivalent. A model in which both intercepts and slopes are random can be specified as shown below:

$$y_{ii} = \pi_{0i} + \pi_{1i}t_{ii} + e_{ii} \quad (15)$$

$$\begin{aligned} \pi_{0i} &= \beta_{00} + u_{0i} \\ \pi_{1i} &= \beta_{10} + u_{1i} \end{aligned} \quad (16)$$

Notably, all of the parameters in this model are interpreted exactly the same as described in the first unconditional means model with random intercepts and fixed slopes. The only difference in this model is that a random effect is specified (u_{1i}). The u_{1i} random

effect is the residual value for each slope and thus is interpreted as the difference in individual i 's slope from the overall slope. The specification of u_{ji} allows slopes to vary and thus a τ_{11} parameter can be estimated. The τ_{11} parameter indicates the variation in slopes, just as τ_{00} indicates the variation in intercepts. Examination of the variation for both of the level two residual estimates provides information regarding the spread of the individual slopes and intercepts about the overall mean slope and intercept. The covariance between the individual slopes and intercepts, τ_{10} is also examined and indicates the relationship between the two parameters. For example, this parameter can indicate whether individuals who start high on sense of identity tend to increase at a higher rate than those who start lower.

The resulting \mathbf{V} matrix from a random coefficient model with both random intercepts and slopes does not correspond to any of the residual covariance matrix forms described previously. The \mathbf{G} , \mathbf{R} and resulting \mathbf{V} matrices for this model are shown below in Figure 8.

\mathbf{G}	\mathbf{R}	$\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$
$\begin{bmatrix} \tau_{00} & \\ \tau_{01} & \tau_{11} \end{bmatrix}$	$\begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$	$\begin{bmatrix} \tau_{00} + \sigma^2 & & \\ \tau_{00} + \tau_{01} & \tau_{00} + 2\tau_{01} + \tau_{11} + \sigma^2 & \\ \tau_{00} + 2\tau_{01} & \tau_{00} + 3\tau_{01} + 2\tau_{11} & \tau_{00} + 4\tau_{01} + 4\tau_{11} + \sigma^2 \end{bmatrix}$

Figure 8. Combination of \mathbf{G} and \mathbf{R} to form \mathbf{V} in the random intercept, random slope model⁵

Note how the resulting \mathbf{V} structure of this model allows both variances and covariances to differ. Comparing the random coefficient model with random intercepts and slopes to traditional models, two advantages are clear. First, this model allows

⁵ \mathbf{Z} is a design matrix indicating which effects are random. In this model \mathbf{Z} is a 3x2 vector consisting of a column of ones and a column with values of time, assumed here to be 0, 1, and 2. There are two columns in this matrix because both intercepts and slopes are allowed to randomly vary.

growth trajectories to vary across individuals and provides parameters (τ_{00} , τ_{11}) that capture this variation and are easily interpretable. Second, the resulting residual covariance matrix of this model takes on a form less restrictive than some forms assumed by traditional models and perhaps more in line with the observed covariance matrix.

CHAPTER III: Method

Overview of Analyses

As previously noted, the purpose of this study is twofold. The first purpose of this study is to compare and contrast traditional and modern techniques for analyzing longitudinal data. Specifically, it is of interest to consider the different conclusions researchers may make based on the results from different techniques. Parts A and B of this study, which are described in detail later in the chapter, address this specific purpose. Specifically, repeated measures ANOVA, MANOVA, ACS Modeling, and MLM are compared in terms of overall model fit, specific parameters, and substantive conclusions using data collected on a sense of identity scale from college students. However, it is difficult to directly compare across these methods because the traditional and modern techniques differ in the type of data permitted in analysis as well as the procedures used to analyze the data. Recall that both repeated measures ANOVA and MANOVA require a Type I dataset. As will be seen in the following sections, the data collected for this study is a Type III dataset, a data structure that can only be handled by MLM. Thus, in order to compare between the different techniques in Parts A and B of this study, data was listwise deleted to form a complete dataset. In addition to complete data, Type I data also assumes that all individuals have the same data collection schedule whereas Type III data allows for different data collection schedules. For this reason, the data was treated in Parts A and B as if the collection schedule for each participant was the same. Thus, time was coded with a 0 for initial measurement occasion, a 1 for the second measurement occasion, and a 2 for the third measurement occasion.

Direct comparison between traditional and modern techniques is also difficult due to differences in the procedures commonly used to estimate the models. For this study, the focus was on the use of SAS programs to estimate all models. Researchers most often use PROC GLM when estimating repeated measures ANOVA or MANOVA. However, in order to estimate ACS models or MLM models, PROC MIXED must be employed. Notably, the more traditional techniques *can* be estimated using PROC MIXED; it is simply less common than estimation using PROC GLM. Part A of the study was conducted in order to demonstrate how PROC MIXED can be used to estimate repeated measures ANOVA or MANOVA models and to convey how the results obtained using PROC MIXED are essentially the same as those obtained using PROC GLM.

After establishing agreement between PROC MIXED and PROC GLM, only PROC MIXED was employed for the remaining analyses. In Part B of the study, PROC MIXED was used to estimate several ACS models with varying residual covariance structures, including ACS models most similar to a repeated measure ANOVA and MANOVA. All models were compared to one another to determine the similarities and differences in results when employing different residual covariance structures. In addition, two multilevel models were estimated, compared to one another, and compared to the ACS models to determine similarities and differences in results with different methodology.

In comparing PROC GLM and PROC MIXED for Part A, time was treated as categorical since conventional ANOVA procedures treat time as a categorical predictor. However, in order to ease comparison among models in Part B, time was treated as a continuous predictor. Although the results from these analyses in Part B allow for the

championing of the best fitting model and the most appropriate technique, it is essential to remember that the data was manipulated to resemble the Type I dataset necessary for traditional techniques. Thus for Part C of the analyses, MLM was used to examine change in sense of identity scores over time using the data in its unaltered Type III form. In Part C, the unconditional means model was used to answer the following two research questions:

1. Across measurement occasions and persons, what is the typical level of sense of identity in this college student population?
2. How much variability in sense of identity scores is within persons (across time) and between persons?

Given that there was a significant amount of variability in scores within persons, the unconditional growth model was used in Part C to answer the remaining research questions:

3. Overall, what level of sense of identity do college students have upon entering college?
4. Do students entering college seem to have very similar or very different sense of identity scores?
5. Overall, how do students change in their sense of identity scores as they progress through their college career?
6. Do students differ from one another in how their sense of identity scores change over time?
7. Is there a relationship between students' initial scores of sense of identity and how they change over time?

Participants and Procedure

Sense of Identity scores were collected at three time points for a sample of 9,180 students at a midsized, southeastern university. Not all students provided data at all three time points and the data collection scheme differed across students, resulting in Type III data. Data collection for this study unfolded in two phases. Phase 1 of data collection occurred on university-wide Assessment days. Students still active at the university in fall 2011 who had provided complete data on the Sense of Identity scale on an Assessment Day taking place between fall 2008 and fall 2011 served as the sample in Phase 1 of data collection. In the second phase, the 9180 students resulting from Phase 1 of data collection were emailed in late fall 2011 or early Spring 2012 and asked to provide responses yet again to the Sense of Identity scale. The two sections that follow describe participants and procedures for Phases 1 and 2 of data collection.

Phase 1. Each student is required to take part in a university-wide Assessment Day twice throughout their college career. The first measurement occasion (i.e., fall Assessment Day) occurs the Friday before classes start their freshman year and the second (i.e., spring Assessment Day) typically occurs when the student has obtained between 45-70 credit hours (usually spring semester of their sophomore year). Students who yielded complete data on the Sense of Identity scale at either the fall or spring Assessment Days and who were still active at the university as of fall 2011 were included in the analysis.

Because the Sense of Identity scale has been administered for several years, our sample is comprised of several different cohorts as shown in Table 8. Recall that MLM allows for Type III data and Students who yielded complete data on the Sense of Identity

scale at either the fall or spring Assessment Days and who were still active at the university as of Fall 2011 were included in the study. Notably, 878 individuals provided responses solely on spring Assessment Days, but because they had been a part of the college atmosphere for at least one semester, we did not consider this data to be a measure of “initial” sense of identity. For this reason, these participants were not considered in the analysis. Thus the final sample size was 9,180 students. The resulting sample of 9,180 students is comprised of several different cohorts as shown in Table 8 (e.g., students completing the Sense of Identity scale in FA08 were incoming freshmen in fall 2008). Twelve percent of the students started college in 2008, 14% in 2009, 35% in 2010 and 39% in 2011. Note that 6,957 participants only provided an initial response. Specifically, 350 provided a response in fall 2008, 407 in fall 2009, 2,990 in fall 2010, and 3,210 in fall 2011.

Phase 2. Email addresses from all 9,180 participants were obtained. After obtaining IRB approval, all 5,601 students who completed the scale before fall 2011 were emailed during November 2011 to collect a second or third time point (see Appendix A). The 3,579 students who responded in fall 2011 were also emailed but were sent a slightly different email requesting participation. The only difference between the two emails is that emails to students from the fall 2011 cohort indicated that they would be contacted again in an attempt to obtain another time point (see Appendix B). The fall 2011 students were emailed again in spring 2012 requesting their participation in an attempt to gather more time points (see Appendix C). If students agreed to participate they clicked on a link which took them to a consent form (see Appendix D) in Qualtrics, a web-based survey provider. After providing consent, participants were taken to a page on which they were

asked to respond to the 8 items and provide an email address used to match students current responses to their previous responses (see Appendix E). Students were not offered incentives for responding to the survey.

Because data collection for Phase 2 was done through email, most individuals have a different data collection schedule. Of the 9,180 students emailed, 806 (9%) completed the Sense of Identity scale in Phase 2 (see Table 1). After Phase 2 of data collection, 6,957 (75.78%) participants had complete data for only one measurement occasion, 2,007 (21.86%) participants had complete data for two measurement occasions, and 216 (2.35%) participants had complete data for three measurement occasions.

Measure

The Sense of Identity scale, developed by Lounsbury and Gibson (2011) was employed to measure student sense of identity. As previously discussed, this scale was developed to provide insight as to an individual's day-to-day sense of self. The scale was administered as a part of a general attitudes packet on Assessment Day and as a web-based survey when the students were emailed. Students were asked to respond to the Sense of Identity scale using a 5-point Likert rating from 1 (strongly disagree) to 5 (strongly agree). Responses that fell outside of this range were recoded as missing. For individuals who had complete data on all 8 items, responses were summed to create a total score for Sense of Identity.

Part A

In order to compare results between PROC GLM and PROC MIXED, data were first listwise deleted and treated as if each participant had the same data collection schedule.

Table 8
Summary of Sense of Identity Data Collection

Number of Waves	Phase 1							Phase 2		n	%	%
	FA08	SP09	FA09	SP10	FA10	SP11	FA11	Nov-11	Jan-12			
1	X									350	3.81	75.78
			X							407	4.43	
					X					2990	32.57	
							X			3210	34.97	
2	X	X								7	0.08	21.86
	X			X						618	6.73	
	X					X				46	0.50	
	X							X		23	0.25	
			X	X						2	0.02	
			X			X				732	7.97	
			X					X		28	0.31	
					X	X				11	0.12	
					X		X			1	0.01	
					X			X		237	2.58	
3							X	X		302	3.29	2.35
	X	X						X		2	0.02	
	X			X				X		63	0.69	
	X					X		X		2	0.02	
			X	X				X		1	0.01	
			X			X		X		81	0.88	
Overall					X	X		X	X	66	0.72	100
										9180	100	

Listwise deletion reduced the sample size from 9,180 participants to 216 participants. For these analyses, time was treated as categorical to most closely reflect traditional repeated measures ANOVA and MANOVA. Repeated measures ANOVA was first estimated using PROC GLM. Notably, a separate MANOVA analysis was unnecessary because PROC GLM provides both univariate ANOVA and MANOVA results when a repeated measures univariate ANOVA is estimated. A repeated measures ANOVA was then estimated in PROC MIXED by use of a Type H residual covariance matrix. In order to estimate a MANOVA model, PROC MIXED with an unstructured residual covariance matrix was estimated. The degrees of freedom, F and p -values were compared between the PROC GLM and PROC MIXED results for both the repeated measures ANOVA and MANOVA models.

Part B

The Type I dataset used in Part A was also used in Part B. In contrast to the analyses in Part A, the variable “time” was treated as a continuous predictor in Part B. Six ACS models with differing residual covariance matrices were fit to the Sense of Identity data. Specifically, the compound symmetric, Type H, toeplitz, homogeneous autoregressive, heterogeneous autoregressive, and unstructured residual covariance matrices were estimated. Notably, the Type H and unstructured residual covariance matrices correspond closely with the repeated measures ANOVA and MANOVA models, respectively, in Part A, with the exception of time now being treated as a continuous rather than categorical predictor. All six ACS models were compared to one another based on their model-data fit, fixed effect parameter estimates, and residual covariance matrices. In order to compare the model data fit among the ACS models, the deviance,

AIC and BIC were examined. Smaller values of all three fit statistics indicate better model data fit. Because all of the models are nested within the unstructured residual covariance matrix likelihood ratio tests comparing the compound symmetric, Type H, toeplitz, homogeneous autoregressive, and heterogeneous autoregressive residual covariance matrices with the unstructured residual covariance matrix were performed. By definition the most complex model (the unstructured model) will have the best model-data fit. Non-significant values for the likelihood ratio test, however, indicated whether more parsimonious models fit as well, or not significantly worse than, the unstructured model. In comparing fixed effect parameter estimates (e.g., intercept and slope), the estimates themselves as well as their standard errors were examined for each model. Because each parameter in the unstructured model is freely estimated, comparison of its residual covariance matrix to others provided insight as to how well the more parsimonious models performed.

In addition to the ACS models, two multilevel models were estimated. The first model allowed intercepts to vary but constrained slopes to be fixed. The second model allowed intercepts and slopes to randomly vary. The two multilevel models were compared to one another and the ACS models in terms of model-data fit, fixed effects parameter estimates and their standard errors, and residual covariance matrices. Because the MLM and ACS models are not nested, the AIC and BIC were used to make comparisons among models.

Part C

The unconditional means model, as presented in Equations 8 and 9 in Chapter IIB, was used to answer the first two research questions of Part C. The first research question inquired as to the typical level of students' sense of identity across time. β_{00} was

examined to identify the overall mean of student Sense of Identity scores. Specifically, the value associated with this parameter provided us with an estimate of typical sense of identity scores across students and time. For this parameter, the actual value was of interest, rather than the significance test associated with the value.

The second research question involved variability in sense of identity scores both within persons and between persons. In order to examine within-person variability or variability across time, the σ^2 parameter was examined. Notably, because variance is the sum of the squared deviations from the mean and thus is difficult to interpret, we changed this parameter into a standard deviation by taking its square root. If there was little variability within persons, it would have indicated that the addition of a time variable may not be necessary. In other words, if there was little variability to explain across time in the first place, the addition of time in the model would not have been beneficial. To examine the between-person variability, the τ_{00} parameter was examined. Again, this parameter was changed into a standard deviation in order to make it more interpretable. The variability between individuals was particularly interesting and important because if there was not a significant amount of variability between persons, we would not have needed to use multilevel modeling. Thus, because u_{0i} takes into account the person effect and accounts for the dependency of scores from each person, it is only necessary when there is variability between persons. If this term was non-significant, it would indicate that the dependency of the scores was not an issue and we would have been able to use traditional regression techniques to analyze the data. In examining the τ_{00} parameter, we were particularly interested in the ratio of between-person variability, τ_{00} , to total

variability, $\tau_{00} + \sigma^2$. This value is the intraclass correlation coefficient (ICC), which represents the proportion of total variability that is between individuals.

Given that there was significant within-person variability to be explained in the unconditional means model, the unconditional growth model, as shown in Equations 14 and 15, was used to answer the last five research questions described in the beginning of Chapter 3. The third research question, inquired as to students' typical sense of identity scores as they enter college. The β_{00} parameter was examined to inform us as to the typical students' score when entering college. Again the β_{00} parameter in the unconditional growth model was slightly different than the unconditional means model. In the unconditional growth model it was the students' typical sense of identity score *when time is equal to zero*. Again, the actual value was more of interest than the significance test associated with this value.

The fourth research question involved the examination of whether or not students entering college tend to have similar or different scores in sense of identity. In order to investigate this, the τ_{00} was examined to inform us as to the spread of individual intercepts. If τ_{00} was significantly different than zero, it indicated that individuals enter college with different levels of sense of identity. As with the terms in the unconditional means model, the standard deviation was used to aid in interpretations. The standard deviation was also used to help determine a plausible values range of individual intercepts. The 95% plausible value range was obtained by multiplying the standard deviation by 1.96 and adding and subtracting the resulting value from the mean. This range provided an indication as to the range in sense of identity scores at the initial time point.

The fifth research question examined how, overall, students' change in their sense of identity scores as they progressed through college. For this question, we examined the overall slope, or β_{10} , parameter. This estimate informed us as to the overall trajectory of scores across individuals. We were interested in both the value and the significance of this parameter; that is, whether the slope was significantly different than zero. Thus, a significant overall slope would indicate that students' sense of identity scores seem to be changing over time. σ^2 indicates the amount of variability in sense of identity scores that cannot be explained by time and can be compared to the σ^2 from the unconditional means model. This comparison allowed us to examine whether or not time was a practically significant predictor of sense of identity scores. Specifically, the pseudo R^2 statistic discussed in Chapter IIB was used to examine what proportion of variance could be explained with the variable time.

The sixth research question for this study asked whether students change differentially in their sense of identity scores over time. Examination of τ_{11} helped to inform us as to whether or not there was variability in individual slopes. Specifically, the significance of τ_{11} was examined. If there was significant variability in the slopes, that would indicate that individuals seem to change differently in sense of identity over time. Again, the standard deviation was used to help with interpretations. Again, about two standard deviations were added and subtracted to the mean to form a plausible values range for student change in sense of identity over time in the population.

The seventh research question inquired as to the relationship, or lack of relationship, between individuals' initial sense of identity scores and how their scores changed over time. For example, examination of this research question helped to inform

us as to whether individuals who started high on sense of identity changed positively, negatively, or not at all over time. The examination of the covariance between the u_{0i} and the u_{1i} terms informed us of this relationship. A significant covariance indicates a significant relationship between how students start out on sense of identity and how they change over time.

Given the varied data collection schedules for the participants, it is important to describe how time was coded in the models. Time was coded as the number of days since the initial measurement occasion for an individual. Because the fall Assessment Day (Phase 1) occurred the Friday before students started classes at the university, it was treated as the initial measurement occasion and time was coded as zero. Time for second measurement occasion was coded as the number of days elapsing between this occasion and the initial measurement occasion. Thus, for a student tested in FA08 and SP10, time at the second measurement occasion was coded as 543, which is the number of days elapsing between their FA08 and SP10 assessments.

CHAPTER IV: Results

Part A

The purpose of this chapter is to demonstrate the similarities and differences between equivalent models estimated using PROC MIXED and PROC GLM in SAS. As previously stated, researchers most often use PROC GLM when estimating repeated measures ANOVA or MANOVA. In order to estimate more modern models, however, PROC MIXED must be employed. Prior to comparing traditional techniques with more modern techniques in the subsequent section where PROC MIXED was used to estimate all models, it is important to demonstrate that the results obtained with PROC MIXED and those obtained with PROC GLM are essentially equivalent.

Comparing PROC GLM and PROC MIXED

Although both PROC MIXED and PROC GLM can be used to analyze repeated measures data, there are differences between the two procedures. The first difference between these two procedures involves the type of data permitted. PROC GLM only allows for Type I data to be analyzed, whereas PROC MIXED allows for Type II (with ACS models) and Type III data (with MLM models). If the dataset is not a Type I dataset, listwise deletion must be used to force it to resemble a Type I dataset. Thus, PROC MIXED is the more flexible procedure of the two with regard to the type of data that can be analyzed. In addition, the two procedures differ in the types of residual covariance structures that can be specified. PROC GLM only uses Type H and unstructured residual covariance matrices for univariate repeated measures ANOVA and MANOVA, respectively. PROC MIXED, on the other hand, allows for specification of Type H and unstructured residual covariance matrices as well as several other residual

covariance matrices. As previously discussed, specification of an appropriate residual covariance structure is essential to ensure trustworthy results. Another difference between the two procedures is the type of estimation employed. By default, PROC GLM uses method of moments estimation, whereas PROC MIXED uses restricted maximum likelihood estimation (Wolfinger & Chang, 1995). The use of restricted maximum likelihood estimation is what allows PROC MIXED to use Type III data. Notably, when there is no missing data, the results from PROC GLM and PROC MIXED should be equivalent (Wolfinger & Chang, 1995).

Applied Example

In order to demonstrate the equality of the two procedures when there is no missing data (as with Type I data), repeated measures univariate ANOVA and MANOVA were estimated using the Sense of Identity data. Because the dataset collected was a Type III dataset, listwise deletion was employed to create a Type I dataset.

Descriptive statistics for the Type I dataset are shown in Table 9.

Table 9

Descriptive Statistics for the Type I Dataset (N = 216)

Time	<i>M</i>	<i>SD</i>	<i>N</i>	V
1	32.70	4.39	216	$\begin{bmatrix} 19.29 & .52 & .39 \\ 11.14 & 24.15 & .57 \\ 8.48 & 13.93 & 24.69 \end{bmatrix}$
2	32.34	4.91	216	
3	32.87	4.97	216	

In all analyses, measurement occasions were treated as categorical variables and it is wrongly assumed that each participant has equal distance between measurement occasions. First, a repeated measures univariate ANOVA was estimated in PROC GLM. The output provided for a repeated measures univariate ANOVA includes both ANOVA and MANOVA results when estimated using PROC GLM. Next, a repeated measures

univariate ANOVA was estimated with PROC MIXED and a Type H residual covariance matrix. The Type H residual covariance matrix was employed because this matrix is used in estimating the repeated measures univariate ANOVA in PROC GLM (Wolfinger & Chang, 1995). Subsequently, a MANOVA was estimated with PROC MIXED and an unstructured residual covariance matrix. Again, the unstructured covariance matrix was employed because MANOVA assumes an unstructured residual covariance matrix. The results from all analyses are provided in Table 10. As expected, the results for the repeated measures univariate ANOVA and the PROC MIXED analysis with a Type H residual covariance structure are exactly the same. As well, the results for MANOVA estimated with PROC GLM and those estimated using PROC MIXED with an unstructured residual matrix are the same⁶.

Table 10
Comparison of PROC GLM and PROC MIXED

Model	Number of Parameters	Model Fit			<i>df</i>	<i>F</i>	<i>p</i>
		Deviance	AIC	BIC			
Univariate ANOVA							
PROC GLM	--	--	--	--	2, 430	1.39	0.2502
PROC MIXED (Type H)	7	3708.1	3716.1	3729.6	2, 430	1.39	0.2502
MANOVA							
PROC GLM	--	--	--	--	2, 214	1.63	0.1985
PROC MIXED (Unstructured)	9	3701.6	3713.8	3733.9	2, 214	1.63	0.1985

Regardless of whether repeated measures ANOVA or MANOVA are employed to examine change over time, the same substantive conclusions would be made.

Specifically, the null hypothesis ($\mu_1=\mu_2=\mu_3$) would fail to be rejected, suggesting that there are no differences among the three time points.

⁶ Results corresponding to the Hotelling-Lawling Trace statistic were used for PROC GLM. Results for PROC MIXED corresponded to results when the Hotelling-Lawling Trace option was requested.

It is, of course, important to use the most appropriate method of data analysis regardless of the results. Thus consideration should be given as to whether repeated measures ANOVA or MANOVA should be used. To aid in this decision, the sphericity tests should be used. If the assumption of sphericity has not been violated, MANOVA would be underpowered and thus repeated measures ANOVA would be most appropriate. However, if sphericity is not met, as in this situation, MANOVA would be the most appropriate method of data analysis.

Although, the value of epsilon was very close to 1 ($\epsilon_{H-F} = .98$) suggesting that sphericity was likely not an issue, Mauchly's test of sphericity was significant, indicating that sphericity had been violated ($\chi^2(2) = 6.44, p = .04$). The results of Mauchly's test can be obtained in PROC MIXED via a likelihood ratio test comparing the Type H model with the unstructured model (Wolfinger & Chang, 1995). Note that the difference between the deviance of the Type H and unstructured model is within rounding of the test statistic for Mauchly's test in PROC GLM (approximately 6.5). The χ^2 of 6.5 with 2 degrees of freedom produces the same .04 p-value as shown in Mauchly's test above. The assumption of sphericity appears to be violated in the Sense of Identity data based on Mauchly's test; however, the epsilon value suggests that it is only a slight violation of sphericity. Typically, when sphericity has been violated, a MANOVA is used to analyze the data.

As previously stated, PROC MIXED must be used with more modern techniques such as ACS Modeling and MLM in order to specify different residual covariance matrices. Because the results from PROC GLM and PROC MIXED were essentially

equivalent when estimating the traditional techniques, PROC MIXED will henceforth be used for all analyses.

Part B

The purpose of the current section is to compare traditional techniques for analyzing longitudinal data with more modern techniques. Specifically, several ACS models treating time as a continuous variable are estimated using a multitude of residual covariance structures, including those that correspond to traditional analyses (e.g., Type H, unstructured). Additionally, two MLM models are estimated. The first is a model in which intercepts are random and slopes are fixed whereas the second is a model in which both intercepts and slopes are random. The ACS models are first compared to each other and subsequently are compared to the MLM models to determine which models best fit the data. Recall that both ACS models and MLM can handle missing data, but ACS models are more restrictive because they assume that individuals have the same schedule of measurement. Thus, the data were, again, wrongly assumed to be Type I.

ACS Models. As previously stated, several ACS models with varying residual covariance structures were estimated. Specifically, six residual covariance structures were examined: compound symmetry, Type H, toeplitz, homogeneous autoregressive, heterogeneous autoregressive, and unstructured. The results for all six models are presented in Tables 11, 12, and 13. Specifically, model fit estimates are presented in Table 11, fixed effects estimates are presented in Table 12, and estimated residual covariance matrices are presented in Table 13. The information from all three tables can be used together to help determine the most adequate residual covariance structure for the data.

First, fit of models relative to one another is examined. The information in Table 11 can be used to help determine the model that produced the best model-data fit. Specifically, the deviance, AIC, and BIC can be compared among models. Notably, smaller values for all three indices are more desirable. Deviances can be compared for nested models using a likelihood ratio test. Thus, because the compound symmetric, Type H, homogeneous autoregressive, heterogeneous autoregressive and toeplitz residual covariance matrices are all nested within the unstructured residual covariance matrix, all models can be compared to the unstructured residual covariance matrix. The unstructured residual covariance matrix is the most complex and by definition will have the best model-data fit. The likelihood ratio test, however, provides information as to whether other residual covariance matrices do not fit significantly worse than the unstructured residual covariance matrix. As shown in Table 11, the toeplitz, homogeneous autoregressive, and heterogeneous autoregressive residual covariance matrices did not fit significantly worse than the unstructured residual covariance matrix. This suggests that any of these three, more parsimonious, models would be adequate for employment in analyzing this data. In addition, the AIC value for the toeplitz residual covariance matrix and the BIC value for the homogeneous autoregressive are the most desirable fit statistics among all six models.

Table 11
Comparing Models

Model	Number of Parameters	Deviance	AIC	BIC	Likelihood Ratio Test		
					$\Delta\chi^2$	Δdf	<i>p</i>
ACS Modeling							
Compound Symmetry	4	3719.7	3723.7	3730.4	14.5 ^a	4	0.0059
Type H	6	3711.3	3719.3	3732.8	6.1 ^a	2	0.0473
Toeplitz	5	3710.3	3716.3	3726.4	5.1 ^a	3	0.1646
Homogeneous Autoregressive	4	3714.0	3718.0	3724.7	8.8 ^a	4	0.0663
Heterogeneous Autoregressive	6	3710.0	3718.0	3731.5	4.8 ^a	2	0.0907
Unstructured	8	3705.2	3717.2	3737.4	--	--	--
Multilevel Modeling							
Model 1 (Random Intercepts, Fixed Slopes)	4	3719.7	3723.7	3730.4	14.5	4	0.0059
Model 2 (Random Intercepts, Random Slopes)	6	3709.6	3717.6	3731.1	4.4	2	0.1108

^a Models compared with the unstructured ACS model

Note. N = 216

Information in Table 12 can help compare the estimated residual covariance matrices of the various models to one another. The unstructured residual covariance structure can, again, be used as a comparison for all of the other residual covariance matrices. This is due to the fact that all parameters in the unstructured residual covariance matrix are freely estimated. Note that in the unstructured matrix, variances increase over time, and adjacent time points have a stronger relationship than non-adjacent time points.

Table 12
Comparing Random Effects Parameters for ACS Models

Structure	Matrix	Parameter	Estimate	Complete
Compound Symmetry	$\begin{bmatrix} \tau_{00} + \sigma^2 & \tau_{00} & \tau_{00} \\ \tau_{00} & \tau_{00} + \sigma^2 & \tau_{00} \\ \tau_{00} & \tau_{00} & \tau_{00} + \sigma^2 \end{bmatrix}$	τ_{00}	11.17	$\begin{bmatrix} 22.74 & .49 & .49 \\ 11.17 & 22.74 & .49 \\ 11.17 & 11.17 & 22.74 \end{bmatrix}$
		σ^2	11.56	
Type H	$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda & \frac{\sigma_1^2 + \sigma_3^2}{2} - \lambda \\ \frac{\sigma_2^2 + \sigma_1^2}{2} - \lambda & \sigma_2^2 & \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda \\ \frac{\sigma_3^2 + \sigma_1^2}{2} - \lambda & \frac{\sigma_3^2 + \sigma_2^2}{2} - \lambda & \sigma_3^2 \end{bmatrix}$	σ_1^2	18.76	$\begin{bmatrix} 18.76 & .49 & .45 \\ 10.86 & 26.09 & .53 \\ 9.53 & 13.19 & 23.42 \end{bmatrix}$
		σ_2^2	26.09	
		σ_3^2	23.42	
		λ	11.56	
Toeplitz	$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}$	σ^2	22.64	$\begin{bmatrix} 22.64 & .54 & .38 \\ 12.14 & 22.64 & .54 \\ 8.71 & 12.14 & 22.64 \end{bmatrix}$
		σ_1	12.14	
		σ_2	8.71	
Homogeneous Autoregressive	$\begin{bmatrix} \sigma^2 & \sigma\rho & \sigma\rho^2 \\ \sigma\rho & \sigma^2 & \sigma\rho \\ \sigma\rho^2 & \sigma\rho & \sigma^2 \end{bmatrix}$	σ^2	22.59	$\begin{bmatrix} 22.59 & .53 & .29 \\ 12.07 & 22.59 & .53 \\ 6.45 & 12.07 & 22.59 \end{bmatrix}$
		ρ	.53	
Heterogeneous Autoregressive	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 \end{bmatrix}$	σ_1^2	19.63	$\begin{bmatrix} 19.63 & .54 & .29 \\ 11.82 & 24.26 & .54 \\ 6.41 & 13.15 & 24.28 \end{bmatrix}$
		σ_2^2	24.26	
		σ_3^2	24.28	
		ρ	.54	
Unstructured	$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$	σ_1^2	19.29	$\begin{bmatrix} 19.29 & .51 & .39 \\ 11.12 & 24.24 & .57 \\ 8.49 & 13.91 & 24.69 \end{bmatrix}$
		σ_2^2	24.24	
		σ_3^2	24.69	
		σ_{12}	11.12	
		σ_{13}	8.49	
		σ_{23}	13.91	

Note. Residual variances are presented on the diagonal, covariances are presented on the bottom off-diagonal, and correlations are presented on the top off-diagonal.

Comparison of these matrices suggests that both the toeplitz and heterogeneous autoregressive residual covariance matrices are similar to those observed in the

unstructured residual covariance matrix. Notably the toeplitz residual covariance matrix seems to best reproduce the relationships, or covariances, between the non-adjacent time points, whereas the heterogeneous autoregressive residual covariance matrix seems to best reproduce the residual variances across measurement occasions. Overall, it seems that the heterogeneous autoregressive model does the best job reproducing the residual variances and covariances, however, this model is more complex than the toeplitz model.

Often models are chosen based on which model has the lowest information criteria. In this study, both toeplitz and homogeneous autoregressive have low information criteria making it difficult to choose between the two models. Recall that the importance of choosing the most appropriate residual covariance matrix is to ensure that the inferences regarding the fixed effects of the model are accurate. Thus to aid in deciding between models, the fixed effects can be examined to determine whether choosing one model over the other would lead to different inferences about the fixed effects.

In examining the fixed effects parameters for all six models in Table 13, it is evident that all six models produce extremely similar estimates for the overall intercept and slope. The overall sense of identity score at the initial time point was estimated to be about 32.5 and the slope was estimated to be about 0.08, indicating that sense of identity increases by 0.08 points for each one unit increase in time. Recall that time was wrongly coded such that the distance between measurement occasions was equal (0, 1, and 2). As noted by Singer and Willett (2003), choice of the residual covariance matrix may not influence parameter estimates, but it can affect their standard errors and therefore inferences made about the significance of parameters. In all models, the intercept was

found to be significantly different than zero and the slope was found not to be significantly different than zero. Additionally, standard errors for each parameter estimate were very similar in magnitude across models.

Table 13
Comparing Fixed Effects Parameters

Model	Intercept	Intercept SE	Slope	Slope SE
ACS Modeling				
Compound Symmetry	32.47*	0.42	0.08	0.16
Type H	32.54*	0.39	0.08	0.16
Toeplitz	32.51*	0.44	0.08	0.18
Homogeneous Autoregressive	32.53*	0.46	0.08	0.19
Heterogeneous Autoregressive	32.60*	0.43	0.07	0.19
Unstructured	32.53*	0.41	0.08	0.18
MLM				
Random Intercepts, Fixed Slopes	32.47*	0.42	0.08	0.16
Random Intercepts, Random Slopes	32.47*	0.41	0.08	0.18

* $p < .01$

Note. N = 216

Multilevel Models. As previously stated, two multilevel models were estimated. Model 1 allowed intercepts to randomly vary, but constrained slopes to be equal for all individuals. As shown in Chapter IIB, this model results in a compound symmetric residual covariance matrix and thus should provide identical results when Type I data is used. Model 2 allowed both intercepts and slopes to randomly vary across individuals. The results for both models are presented in Table 11, Table 13, and Table 14. Specifically, model fit estimates are presented in Table 11, fixed effects estimates are presented in Table 13, and the variance components for the random effects estimates are presented in Table 14. Like comparison of the ACS models, the information from all three tables can be used in conjunction to help determine the most adequate model for the

data. Again, the data have been listwise deleted and is wrongly assumed to have equivalent schedules of measurement for all individuals.

Fit statistics in Table 11 were examined to compare model fit of multilevel models to one another as well to the ACS models previously discussed. Comparison of the two multilevel models indicates that the deviance for the Model 2 is smaller, as expected due to the fact that it is more complex than Model 1. Notably, Model 1 is nested within Model 2 and thus a likelihood ratio test was performed to determine whether Model 1 fit significantly worse than Model 2. The results of this test indicated that Model 1 fit significantly worse than Model 2 ($\chi^2(2) = 10.1, p = 0.0064$). In comparing the two multilevel models to the other ACS models, we see that as expected, the results for Model 1 and the compound symmetric ACS model are exactly the same. Comparison of Model 2 to the unstructured model indicates that Model 2 does not fit significantly worse than the unstructured model ($\chi^2(2) = 4.4, p = .1108$). In order to compare Model 2 to the other ACS models, the AIC and BIC must be used because these models are not nested. Comparison of Model 2 to the other ACS models using the AIC indicates that Model 2 fits the data better than the compound symmetry, Type H, homogeneous autoregressive, and heterogeneous autoregressive residual covariance matrices. Using the BIC suggests that Model 2 fits better than the Type H, heterogeneous autoregressive, and unstructured residual covariance matrices. Both indices suggest that the toeplitz ACS model fits the data better than Model 2.

Table 14
Comparing Random Effects Parameters for Multilevel Models

Model	Matrix	Parameter	Estimate	Complete
Model 1 (Random Intercepts, Fixed Slopes)	$\begin{bmatrix} \tau_{00} + \sigma^2 & \tau_{00} & \tau_{00} \\ \tau_{00} & \tau_{00} + \sigma^2 & \tau_{00} \\ \tau_{00} & \tau_{00} & \tau_{00} + \sigma^2 \end{bmatrix}$	τ_{00}	11.56	$\begin{bmatrix} 22.74 & .49 & .49 \\ 11.17 & 22.74 & .49 \\ 11.17 & 11.17 & 22.74 \end{bmatrix}$
		σ^2	11.17	
Model 2 (Random Intercepts, Random Slopes)	$\begin{bmatrix} \tau_{00} + \sigma^2 & & & \\ \tau_{00} + \tau_{01} & \tau_{00} + 2\tau_{01} + \tau_{11} + \sigma^2 & & \\ \tau_{00} + 2\tau_{01} & \tau_{00} + 3\tau_{01} + 2\tau_{11} & \tau_{00} + 4\tau_{01} + 4\tau_{11} + \sigma^2 & \\ & & & \end{bmatrix}$	τ_{00}	14.11	$\begin{bmatrix} 20.66 & .50 & .43 \\ 10.45 & 21.45 & .56 \\ 9.88 & 13.18 & 26.12 \end{bmatrix}$
		τ_{01}	-2.51	
		τ_{11}	1.94	
		σ^2	9.63	

Table 13 contains the fixed effects parameter estimates for all models. Comparison of the standard errors for the intercepts and slopes of the two multilevel models suggests that Model 2 has a more precise estimate for the overall intercept, whereas Model 1 has a more precise estimate for the overall slope. In comparing the multilevel models to the ACS models, Model 2 has a smaller intercept standard error than the compound symmetric, toeplitz, homogeneous autoregressive, and heterogeneous autoregressive models. The Type H model was the only ACS model with an intercept standard error that was smaller than Model 2's intercept standard error. With regard to slope standard errors, Model 1 has a smaller standard error than the toeplitz, homogeneous autoregressive, heterogeneous autoregressive, and unstructured models. As noted when comparing ACS models in the previous section, all of the standard errors are extremely close in magnitude.

The information in Table 12 and Table 14 allows for comparison of the residual covariance matrices of all models. Again, models can be compared to the unstructured residual covariance matrix to gain insight as to how well the model reproduces the data. Examination of the residual covariance matrices for the multilevel models suggests that Model 2 produces a residual covariance matrix that more closely matches the unstructured residual covariance matrix than Model 1. Again, comparison of the multilevel models with the ACS models confirms the notion that Model 1 is equivalent to the compound symmetric model. The heterogeneous autoregressive model is more accurate with regard to residual variances, and the toeplitz model is more accurate with regard to the relationship between non-adjacent time points compared to Model 2.

However, overall Model 2 is reasonably accurate with regard to the residual covariance matrix.

Overall Comments. Based on the results from the ACS and multilevel models presented above, researchers can evaluate the most appropriate model for data analysis. In conjunction, some of the results may contradict each other. For example, the homogeneous autoregressive model does not fit significantly worse than the unstructured model, but also has the highest standard errors for the fixed effects parameter estimates. Nevertheless, most of the results above advocate for the toeplitz ACS model which suggests that residual variances across measurement occasions are equal and that residual covariances between adjacent measurement occasions are equal (but not necessarily systematically larger than residual variances between non-adjacent measurement occasions). Thus, researchers would conclude that the overall intercept for Sense of Identity scores was 32.51 and that for every unit increase in time, (recall that time was coded with a 0 for initial measurement occasion, 1 for second measurement occasion, and 2 for third measurement occasion) Sense of Identity scores increased by .08.

Even though the toeplitz model seems to be the most appropriate model for the data, it is important to recall a key drawback to ACS models, which is their sole focus on overall change. Note that the toeplitz ACS model does not provide any information as to the individual variation in intercepts or slopes. Model 2 of the multilevel models on the other hand provides information about the overall intercept and slope, as well as an estimate of individual intercept variation (τ_{00}), individual slope variation (τ_{11}), and the relationship between individual slopes and intercepts (τ_{01}). For the Sense of Identity data τ_{00} was 14.11. This value can be used to create a plausible value range for the intercept.

Thus, 95% of the intercepts range between 25.11 and 39.83. The same process can be used with τ_{11} . The 1.94 variance can be used to demonstrate that 95% of the slopes fall between -2.65 and 2.81. Note that even though there is no change in slopes overall, there is variability in individual slopes. Specifically, the plausible value range includes positive and negative slopes, which suggests that some individuals are increasing whereas others are decreasing in Sense of Identity scores over time. In addition to intercept and slope variances, the covariance of -2.51 (or correlation of -.48) between the intercept and slope is provided. Thus, there seems to be a negative relationship between how individuals start on Sense of Identity and how they change over time. Thus, multilevel modeling provides much richer information about individual differences in change over time compared to ACS models.

Of course, a serious weakness of the MLM and ACS models shown here is the substantial manipulation of the original data that had to occur for their use. Recall that listwise deletion was used to force the original dataset into the form of a Type I dataset. Thus, the original sample of 9,180 participants was reduced to 216 participants, greatly reducing the power of the analyses and, depending on the type of missing data, biasing parameter estimates. In addition, the data was treated as if the data collection schedules for all 216 remaining participants were the same whereas, in reality, most individuals had differing data collection schedules. Due to the nature of the original dataset, estimating models (e.g., ACS models) that assume Type I or Type II data, would be inappropriate. Given that MLM is the only technique that allows for Type III data, the next set of analyses was conducted to most appropriately analyze the Sense of Identity data.

Part C

Prior to fitting the multilevel models, a graph displaying the trajectories of a random sample of 25 students with data at all 3 time points was created to obtain a sense of how sense of identity changes over time and individual variation in change. As can be seen from Figure 9, individuals differ in how they start off and how they change in sense of identity scores over time. In addition, the graph displays the fact that some individuals have very different schedules of measurement. Some individuals have their first, second, and third measurement occasions within a 200 day period, whereas other individuals have all three measurement occasions spread across 1200 days. In order to model differing schedules, time in the multilevel models was coded as number of days since initial measurement occasion.

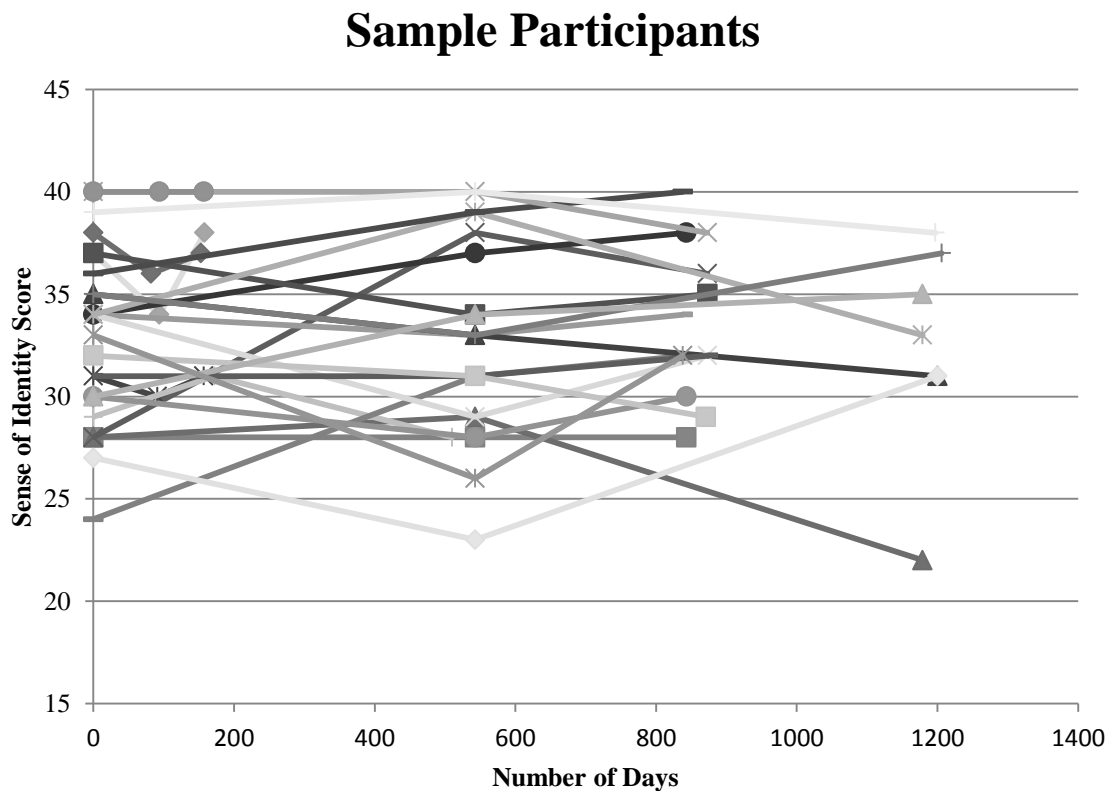


Figure 9. Measurement schedule and trajectories for 25 participants

Two models, the unconditional means model, and the unconditional growth model, were fit to the data to answer the seven research questions presented in Chapter III. Table 15 presents the results from the Part C analyses. The unconditional means model, or intercept-only model, was the first to be estimated in Part C. This model answers the first two research questions presented in Chapter III regarding the typical level of Sense of Identity and the variability within- and between-persons. The estimate for β_{00} in the unconditional means model indicates the typical level of Sense of Identity across individuals and across time points. Thus, overall, students tend to have Sense of Identity scores of about 32.36. The estimate for σ^2 in the unconditional means model indicates the amount of within-person variability in Sense of Identity scores. Thus the value of 12.02 indicates that scores within individuals tend to deviate from the individual's average Sense of Identity score by about 3.47 points. The estimate for τ_{00} in the unconditional means model indicates the amount of between-subjects variability and is significantly different than zero in the present study. The value of 14.64 suggests that individuals deviate from the overall mean by about 3.83 points. τ_{00} can be used to create a plausible values range for the overall intercept in the unconditional means model. Addition and subtraction of approximately 2 times $\sqrt{\tau_{00}}$ provides a range in which Sense of Identity values are likely to be within the population. Thus, for this study, Sense of Identity scores are likely to range from 24.85 to 39.87. Notably, this is a fairly large range of values and suggests that there is a fair amount of between-person variability.

τ_{00} and σ^2 from the unconditional means model can also be used to calculate the intraclass correlation coefficient (ICC). The ICC indicates the amount of total variability

that is between-individuals. The ICC for this study is .55, indicating that more than half of the variability in Sense of Identity scores is between-individual. This is a large ICC value and supports the idea that person effects should be included in the model.

Table 15
Fixed and Random Effects for the Unconditional Mean and Unconditional Growth Models

Model	Fixed Effects		Random Effects		
	Parameter	Estimate	Parameter	Estimate	<i>p</i>
Unconditional Means Model	β_{00}	32.36 (.05)*	τ_{00}	14.64	<0.001
			σ^2	12.02	---
Unconditional Growth Model	β_{00}	32.33 (.05)*	τ_{00}	14.72	<0.001
	β_{10}	.0004 (.00)*	τ_{10}	-.0002	1.0000
			τ_{11}	.0000	1.0000
			σ^2	11.97	---

Note. Values in parentheses are standard errors

* $p < .01$

The unconditional growth model was estimated to answer the remaining research questions presented in Chapter III. Again, fixed effects and variance components for the random effects for the model are presented in Table 15.

The estimate for β_{00} in the unconditional growth model answers the third research question regarding individuals' Sense of Identity scores upon entering college. Recall that the intercept in the unconditional growth model differs from that in the unconditional means model in that it is the overall average Sense of Identity score when time is equal to zero. In this study, the initial measurement occasion for each participant was coded as a zero and thus β_{00} represents the average Sense of Identity score at the initial measurement occasion (for all students this was the beginning of freshman year). Thus, on average, students enter college with a Sense of Identity score of about 32.33.

The estimate for τ_{00} provides information as to how intercepts vary in the population and is significantly different from zero in the current study. Because τ_{00} is a variance, the square root can be taken to aid in interpretation. Thus, students' intercepts tend to vary about the overall intercept by about 3.84 points. Again, τ_{00} can be used to create a 95% plausible values range: intercepts are likely to be between 24.80 and 39.86.

The estimate for β_{10} is used to answer the fifth research question regarding whether or not individuals change in Sense of Identity over time. Recall that the interpretation of this parameter is that for every unit change in time, there is β_{10} change in Sense of Identity. Thus, the interpretation changes with the coding of time. In this study, time was coded as days between measurement occasions. The β_{10} value indicates that for each day there is a .0004 change in Sense of Identity. In order to examine change in Sense of Identity over a longer period of time, a year for example, the slope is simply multiplied by 365. Thus for each year, Sense of Identity increases by .146 points. On a scale that ranges from 8 to 40, this is an extremely small change, suggesting that individuals are not changing in Sense of Identity over time. The slope parameter is significant, $t(2439) = 2.62, p = .009$. However, the significance is likely due to the fact that the sample is very large. In addition to β_{10} , σ^2 can be examined between the unconditional means and unconditional growth model to determine whether or not time was a practically significant predictor of Sense of Identity scores. The difference between σ^2 in the unconditional means model and σ^2 in the unconditional growth model represents the amount of variability in Sense of Identity scores that can be explained by time. Thus, a proportion of variance explained by time to total variability can be

calculated to demonstrate the percent of total variability explained by time. This is also known as the Pseudo R^2 statistic and is 0.42% in this study. In other words, time can only explain less than 1% of the variability in Sense of Identity scores, and thus is not a practically significant predictor.

The estimate for τ_{11} indicates the amount of variability in individual slopes and thus answers the sixth research question presented in Chapter 3. Notably the value of τ_{11} is .0000 (the exact value was .0000000172) indicating that individuals' slopes do not vary. In addition, the estimate of τ_{11} was not significantly different than zero. This indicates that the random effect for the slope parameter is unnecessary and a more parsimonious model would adequately model the data.

Typically, the relationship between the slopes and intercepts can be examined to determine whether the way an individual starts on Sense of Identity is related to how they change over time. Thus researchers could answer questions such as whether or not individuals who start high on Sense of Identity continue to increase, decrease, or stay the same over time. The estimate for τ_{10} provides information about the relationship between slopes and intercepts. Because τ_{10} is a covariance, however, it is difficult to interpret and thus it can be transformed to a correlation. In this study the correlation between slopes and intercepts is -.30. This may seem like a sizeable relationship, but is not significant. It is likely that the seemingly sizeable -.30 relationship is found because in converting a covariance to a correlation the following equation is used.

$$r_{(x,y)} = \frac{\text{cov}(x,y)}{s_x s_y} \quad (17)$$

Note that the square root of the residual intercept and slope variance (making them the intercept and slope standard deviation, denoted by “s”) are multiplied to form the denominator. Because the slope variation was extremely small the denominator was also extremely small, which may have contributed to a seemingly sizable correlation coefficient.

CHAPTER V: Discussion

Recall that this study had two main purposes. First, traditional and modern techniques for analyzing longitudinal data were compared and contrasted using an applied example with Sense of Identity data at three measurement occasions. Subsequently, multilevel modeling was used to examine change in Sense of Identity scores over time.

Often, researchers gather longitudinal data and repeated measures ANOVA or MANOVA are used to examine whether means differ across measurement occasions. Repeated measures ANOVA and MANOVA make several assumptions about the data that must be met in order to obtain trustworthy results. Specifically, both ANOVA and MANOVA assume Type I data. Thus, for the Sense of Identity data, which was collected as a Type III dataset, observations had to be listwise deleted in order to force the data into a Type I form. Listwise deleting data not only biases parameter estimates, but also drastically reduces the power of the analyses. In the Sense of Identity data listwise deletion reduced the sample of 9,180 to 216. Additionally, Type I data assumes that each individual has the same amount of time between measurement occasions. Because the Sense of Identity data is a Type III dataset, each individual has a different schedule of measurement. Table 16 below provides descriptive statistics for the Time variable using the reduced sample of 216 students to demonstrate the average number of days for the second and third time points.

Table 16
*Descriptive Statistics for Time Variable as Represented
 as Number of Days Since Initial Measurement Occasion*

	<i>M</i> (in days)	<i>SD</i>	Min	Max
Time 0	0.00	0.00	0.00	0.00
Time 1	401.03	215.91	74.00	907.00
Time 2	743.79	418.34	150.00	1206.00

Note. N=216

The table above demonstrates the different schedules of measurement for individuals. Note that individuals' second measurement occasion can range from 74 to 907 days from initial measurement occasion. If a repeated measures ANOVA or MANOVA were estimated for the data, researchers would be forced to either treat the data as if each individual had the same schedule of measurement, or delete individuals until a true Type I dataset can be formed. In Part A and B demonstrations, the former alternative was utilized. As shown in Table 16 above, treating the data as if measurement occasions were equivalent across individuals is clearly inappropriate. For example, individuals with 74 and 907 days in between their first and second time points were treated as if they had the same number of days in between time points. Similarly, individuals with 150 to 1206 days in between the first and third time point were treated as if they had the same number of days in between these time points. Thus, treating the data as if measurement occasions were equivalent across individuals is clearly inappropriate. For this reason, the results from Parts A and B are not used to make inferences regarding change in Sense of Identity over time, but to compare traditional with modern techniques, as well as to show how biased results are when treating the data incorrectly.

Part A was used to demonstrate how both repeated measures ANOVA and MANOVA can be estimated using PROC GLM or PROC MIXED in SAS. As illustrated

by the results for Part A, equivalent results are achieved regardless of whether PROC GLM or PROC MIXED is used. Notably, these results are only equivalent because a Type I dataset was used. If the dataset contained missing data, the results from PROC GLM and PROC MIXED would differ because PROC GLM uses listwise deletion whereas PROC MIXED allows all data to be used in analyses.

Although the results from Part A cannot be used to make conclusions about how Sense of Identity changes over time given the manipulation of the data, it is of interest to consider what conclusions a researcher utilizing these methods would make. The results from the repeated measures ANOVA and MANOVA in Part A both indicated that the average Sense of Identity scores do not differ across time. Although the conclusions don't differ between repeated measures ANOVA and MANOVA, a researcher would typically choose to report the results of one method over the other. Conventional researchers would typically examine whether or not the assumption of sphericity was met in order to decide between repeated measures ANOVA and MANOVA. Mauchly's test of sphericity indicated that sphericity had been violated, however epsilon was extremely close to 1 suggesting slight, if any, violation of sphericity. Given the significance of Mauchly's test, it is possible that many researchers would choose MANOVA to analyze mean differences.

As previously mentioned, use of MANOVA when the assumption of sphericity holds reduces the power of the analyses. Because sphericity was only mildly violated as indicated by epsilon, it is possible that MANOVA is underpowered, but given that it provides the same conclusion as repeated measures ANOVA, this does not appear to be the case. Often another drawback of MANOVA is the use of the unstructured residual

covariance matrix. By definition the unstructured residual covariance matrix is the *most* appropriate residual covariance matrix; however, it is also the *least* parsimonious residual covariance matrix. MANOVA requires a large sample size in order to precisely estimate all of the parameters in the model; however, in this case there are only three time points and thus only 3 variances, 3 covariances and 2 fixed effects. Thus, many researchers would feel confident being able to precisely estimate this small number of parameters with a sample of 216.

Some researchers may be torn between repeated measures ANOVA and MANOVA due to the fact that Mauchly's test was significant, but epsilon was extremely close to 1. Thus, in order to estimate models with residual covariance matrices that are more flexible than Type H, and more parsimonious than unstructured, ACS models can be employed. Part B of the analyses estimated models with differing residual covariance matrices to find the most appropriate and the most parsimonious model. For these analyses, time was treated as a continuous variable.

An advantage of ACS modeling is that it allows for the specification of a wide range of residual covariance matrices. Thus, a residual covariance matrix that is both parsimonious and appropriate can be employed, unlike in repeated measures ANOVA and MANOVA. The specification of a parsimonious residual covariance matrix is an advantage because simpler models are often more desirable. Ensuring that an appropriate residual covariance matrix is employed is important because it can affect the inferences made about the fixed effects in the model. In the present study, however, the results from six ACS models differing in residual covariance matrices were extremely similar and all

suggested that the slope parameter was not significant. In other words, all ACS models indicated that Sense of Identity scores did not significantly change in a linear fashion over time. Thus, the same substantive conclusions would be made about change in Sense of Identity over time regardless of which ACS model, including those most similar to repeated measures ANOVA and MANOVA, had been employed. Thus, the advantages of ACS modeling were not realized in Part B of this study.

It is important to note, that although Part B of the study used the manipulated Type I data, ACS models are not limited to Type I data. Because ACS modeling is estimated using PROC MIXED and maximum likelihood estimation, it is much more flexible than repeated measures ANOVA and MANOVA in that it allows the use of Type II data. Recall that the only difference between Type I and Type II data is that Type II data allows for missing data. Notably the allowance of missing data would allow for all 9,180 individuals in the dataset to be used thus providing much more power for the analyses than repeated measures ANOVA or MANOVA. A drawback to ACS models, however, is that measurement schedules need to be the same for all individuals. Again, the Sense of Identity data was a Type III dataset, so in order to estimate ACS models, we had to pretend as if each individual has the same schedule of measurement.

ACS modeling, repeated measures ANOVA, and MANOVA are all similar in that that the focus of all three techniques is on overall change across measurement occasions. Although a linear trend can be specified in repeated measures ANOVA, MANOVA, and ACS modeling, none of these techniques provide parameters that easily allow interpretation of differences in individual change over time. Again, these parameters may

not be of interest for some research questions, but they provide much more information about change over time than solely examining mean change over time.

Multilevel modeling is a much more flexible technique that can be used with longitudinal data. Multilevel modeling offers an advantage over traditional models in that it allows for individual change over time to be examined. Thus, overall differences across time can be examined as well as variability in individual change over time. The fact that multilevel modeling allows for the examination of individual change over time offers much richer and more useful information than the information offered by the traditional techniques.

Multilevel modeling is also advantageous in that it is the only technique out of the four discussed that allows for a Type III dataset. Thus, all individuals can contribute to the analyses and researchers do not have to treat the data as if each individual has the same schedule of measurement. In sum, this method is most appropriate for the type of Sense of Identity data collected for this study. In addition, multilevel modeling also allows the advantage of providing a more flexible residual covariance matrix. Again, constraining parameters in the model can produce familiar residual covariance matrices (e.g., compound symmetry). When parameters are not constrained, however, the combination of the **G** and **R** matrices allows for a flexible residual covariance matrix that is more parsimonious than the unstructured matrix used in MANOVA.

Because multilevel modeling offers significant advantages over the traditional techniques, it was employed to examine change in Sense of Identity over time. The results, overall, suggested that on average, students entered college with moderately high Sense of Identity scores (average of 32.33 on a scale ranging from 8 to 40) and that

scores did not change over time. The results indicated substantial variability among individuals in Sense of Identity scores upon entry to college, with 95% of the intercepts in the population ranging from about 25 to 40, which captures the midpoint of the scale to the highest value. Thus, very few students have low sense of identity coming into college. Although there was substantial variability in individual intercepts, there was no variability in slopes. Thus, the finding that scores do not change over time in the overall sample applies to individuals as well. Given these results, the unconditional means model would be adequate to model the data. There were significant differences between the unconditional means and unconditional growth models due to the significant slope parameter in the unconditional growth model. Recall, however, that the slope parameter was essentially zero and was likely significant due to the large sample size. Because the slope parameter and variation in slopes were both essentially zero, they are not necessary in the model. Notably, the results for the multilevel models in Part B, when treating the data as a Type I dataset, suggested that there was significant slope variation. In turn, researchers who used the manipulated Type I data set would conclude that there is slope variation when, in reality, there is not. This demonstrates the possible consequences of utilizing an altered data set.

Although our conclusions about change in Sense of Identity scores over time from the multilevel model is similar to the repeated measures ANOVA model, multilevel modeling still offers advantages over this traditional technique. Most importantly, multilevel modeling allows for all 9,180 participants to be included in data analysis. Thus, researchers can be confident that the analyses have enough power *and* can be confident that the parameter estimates are not biased due to listwise deletion of missing

data. In addition, multilevel modeling allows for differing schedules of measurement and thus the data did not need to be misleadingly treated as if all individuals were on the same schedule of measurement. Additionally, had multilevel modeling not been employed, researchers would have to *assume* that the slopes were the same across individuals. Multilevel modeling allows researchers to empirically test this assumption and thus specify the most appropriate model.

Past research, presented in Chapter IIA suggested that Sense of Identity would change, overall, over time and that individuals would vary in the way they change in Sense of Identity over time. Because the results did not support our hypotheses it is important to consider explanations as to why Sense of Identity scores did not change over time. Although it is possible that Sense of Identity truly does not change over time, the research presented in Chapter IIA suggests that it is a developmental process and should change as time progresses. Recall that the Sense of Identity scale most closely aligns with Marcia's identity achievement category and that students entered college with a fairly high Sense of Identity. Thus, it is possible that individuals do not fluctuate as much as originally anticipated once identity has been achieved.

It is also possible that Sense of Identity does change over time, but that the Sense of Identity scale does not measure the construct well. Past research indicates that some of the items may need to be removed or omitted (Samonte & Pastor, 2011). If items on the scale are not functioning well, the Sense of Identity scores may not be meaningful. It is also possible that the core sense of self that the Sense of Identity scale aims to measure is too broad to examine changes over time. It may be that more specific parts of identity change over time, as seen in past literature, but that general identity does not fluctuate as

greatly. As such, it may benefit researchers to examine both general identity and identity in specific domains simultaneously over time. In addition, the midpoint of the Sense of Identity scale is labeled “Neutral/Undecided.” It is possible that individuals who do not have a well-developed, strong sense of identity would endorse an “undecided” option rather than the “strongly disagree” option. Thus, different students may not use the response scale the same way and responses at the low *or* middle of the scale would indicate low Sense of Identity depending on how the individual interpreted the scale. If this is the case, scores on the Sense of Identity scale cannot be interpreted in a meaningful way.

Additionally, it is possible that the “treatment” (college) expected to increase Sense of Identity scores is not as influential as originally anticipated. If college is not a treatment that influences identity change, it would not be surprising to see that identity did not change throughout time spent at college. It is important to note, however, that much of the research presented in Chapter IIA suggests that late adolescence and the college years are an ideal time for identity change. It may be useful to extend the time of measurement to examine the years before, during, and after college.

Limitations and Future Directions

There are several limitations that should be noted in this study. First, researchers should thoroughly consider whether or not it is appropriate to model individuals with 74 days between the first and second time point with individuals with 907 days between the first and second time point in the same model. Although multilevel modeling can handle this type of data, researchers may want to consider whether individuals with a shorter distance between measurement occasions may have different slopes than individuals with

a longer distance between measurement occasions. Predictors such as cohort may be added to the model to examine whether individuals from differing cohorts have different slopes. In the future, researchers should consider a more structured data collection schedule. Although it is not necessary that individuals have the exact same schedule of measurement, it may be beneficial to examine change in sense of identity over a semester or over years rather than both at the same time. Examination of Sense of Identity over a longer period of time (e.g., throughout college and after graduation) may also provide more insight as to changes in Sense of Identity throughout early adulthood.

Second, only 3 waves of data were collected from participants. When only three measurement occasions are collected, only a linear model can be fit to the data. If the relationship between sense of identity and time was quadratic or cubic, more measurement occasions would need to be collected before the appropriate relationship could be modeled. Specifically, 4 measurement occasions would be necessary to model a quadratic relationship whereas 5 measurement occasions would be necessary to model a cubic relationship. In future studies, additional measurement occasions would allow for a more accurate model to be estimated. Thus, at least four measurement occasions should be gathered to examine a possible quadratic relationship between sense of identity and time.

Third, the number of participants providing three waves of data ($N = 216$) is drastically smaller than the original sample size ($N = 9,180$). Thus, it is possible that the 216 participants who provided all three waves of data differ from those who only provided one or two waves of data. Relatedly, it may be that individuals who chose to respond to the survey, regardless of whether it was their second or third measurement

occasion, differ than those who did not. Adding predictors, such as conscientiousness may help to predict survey completion. Knowledge of the types of individuals most likely to complete the survey would help inform researchers of the population the results would be most applicable.

Final Conclusions

This thesis presented a strong case in favor of considering more modern methods for analyzing longitudinal data. Specifically, multilevel modeling was argued to be more appropriate when there is missing data and/or when the data is unbalanced on time. Multilevel modeling is the most flexible technique with regard to type of data permitted in analyses and thus allows for the optimal use of information. Additionally, multilevel modeling provides richer, more interpretable information than traditional techniques regarding individual variability in change over time. Thus, researchers interested in individual variation in change over time would greatly benefit by use of multilevel modeling.

The findings from the study suggest that students enter college with a moderately high level of Sense of Identity. Additionally, the results suggest that students do not linearly change in Sense of Identity levels throughout college. Notably, more research, as outlined in the previous section, should be conducted before conclusions are made suggesting that Sense of Identity is stable throughout college.

Appendix A

Dear JMU Student,

My name is Kelli Samonte and I am a 2nd year master's student at JMU in the Psychological Sciences program. My advisor, Dr. Dena Pastor of Graduate Psychology, and I are interested in how JMU students' sense of identity changes over the course of their college career. In order to examine this, we need to measure students' sense of identity on multiple occasions throughout their JMU experience.

You are receiving this email because during Assessment Day at JMU you completed a sense of identity scale. I am hoping that you would be willing to complete this 8-item scale again so that I may examine how our students' sense of identity changes over time here at JMU. **Your participation is completely voluntary and the entire 8-item survey should take you no longer than 5 minutes.**

The link below will direct you to a consent form and subsequently the following 8 items:

1. I have a definite sense of purpose in life.
2. I have a firm sense of who I am.
3. I have a set of basic beliefs and values that guide my actions and decisions.
4. I know what I want out of life.
5. I have a clear set of personal values or moral standards.
6. I don't know where I fit in the world.
7. I have specific personal goals for the future.
8. I have a clear sense of who I want to be when I am an adult.

In order to link your responses with those that you provided on Assessment day, we do request that you provide your JMU email address, but can ensure that this is solely to match your responses. All responses will be kept completely confidential. Your participation is voluntary, but we do hope you choose to participate.

To participate, please use the following link:

****hyperlink****

We thank you in advance for your participation!

Kelli Samonte
2nd Year Master's Student
Psychological Science Program
Department of Graduate Psychology
samontkm@dukes.jmu.edu

Dr. Dena A. Pastor
Associate Professor
Department of Graduate Psychology
pastorda@jmu.edu

Appendix B

Dear JMU Student,

My name is Kelli Samonte and I am a 2nd year master's student at JMU in the Psychological Sciences program. My advisor, Dr. Dena Pastor of Graduate Psychology, and I are interested in how JMU students' sense of identity changes over the course of their college career. In order to examine this, we need to measure students' sense of identity on multiple occasions throughout their JMU experience.

You are receiving this email because during Assessment Day at JMU you completed a sense of identity scale. I am hoping that you would be willing to complete this 8-item scale again so that I may examine how our students' sense of identity changes over time here at JMU. In order to gain a better idea of how sense of identity changes over time you will also receive an email in Spring 2012 asking you to complete this survey one last time. **Your participation is completely voluntary and the entire 8-item survey should take you no longer than 5 minutes.**

The link below will direct you to a consent form and subsequently the following 8 items:

1. I have a definite sense of purpose in life.
2. I have a firm sense of who I am.
3. I have a set of basic beliefs and values that guide my actions and decisions.
4. I know what I want out of life.
5. I have a clear set of personal values or moral standards.
6. I don't know where I fit in the world.
7. I have specific personal goals for the future.
8. I have a clear sense of who I want to be when I am an adult.

In order to link your responses with those that you provided on Assessment day, we do request that you provide your JMU email address, but can ensure that this is solely to match your responses. All responses will be kept completely confidential. Your participation is voluntary, but we do hope you choose to participate.

To participate, please use the following link:

****hyperlink****

We thank you in advance for your participation!

Kelli Samonte
2nd Year Master's Student
Psychological Science Program
Department of Graduate Psychology

samontkm@dukes.jmu.edu

Dr. Dena A. Pastor
Associate Professor
Department of Graduate Psychology
pastorda@jmu.edu

Appendix C

Dear JMU Student,

My name is Kelli Samonte and I am a 2nd year master's student at JMU in the Psychological Sciences program. My advisor, Dr. Dena Pastor of Graduate Psychology, and I are interested in how JMU students' sense of identity changes over the course of their college career. In order to examine this, we need to measure students' sense of identity on multiple occasions throughout their JMU experience.

You received an email in the Fall asking you to complete this survey. The current email is in hopes that you would be willing to complete this 8-item scale again, regardless of whether or not you responded in the Fall. Your participation will help me to examine how our students' sense of identity changes over time here at JMU. **Your participation is completely voluntary and the entire 8-item survey should take you no longer than 5 minutes.**

The link below will direct you to a consent form and subsequently the following 8 items:

1. I have a definite sense of purpose in life.
2. I have a firm sense of who I am.
3. I have a set of basic beliefs and values that guide my actions and decisions.
4. I know what I want out of life.
5. I have a clear set of personal values or moral standards.
6. I don't know where I fit in the world.
7. I have specific personal goals for the future.
8. I have a clear sense of who I want to be when I am an adult.

In order to link your responses with those that you provided on Assessment day, we do request that you provide your JMU email address, but can ensure that this is solely to match your responses. All responses will be kept completely confidential. Your participation is voluntary, but we do hope you choose to participate. Again, even if you did not respond in the Fall your participation now would be greatly appreciated.

To participate, please use the following link:

****hyperlink****

We thank you in advance for your participation!

Kelli Samonte
2nd Year Master's Student
Psychological Science Program
Department of Graduate Psychology

samontkm@dukes.jmu.edu

Dr. Dena A. Pastor
Associate Professor
Department of Graduate Psychology
pastorda@jmu.edu

Appendix D

Consent to Participate in Research

Identification of Investigators & Purpose of Study

You are being asked to participate in a research study conducted by Dr. Dena Pastor (Department of Graduate Psychology) and Kelli Samonte (Department of Graduate Psychology). The purpose of the present study is to examine students' sense of identity over time. Responses collected from this survey will be used to inform researchers as to how students' sense of identity changes and develops throughout their college experience.

Research Procedures

This study consists of an online survey that will be administered to individual participants through *Qualtrics*, an online survey tool. You will be asked to provide answers to a series of items related to your sense of identity. Should you decide to participate in this confidential research you may access the survey by following the web link provided. You will be asked to provide your 9-digit JMU student identification number in order to match your responses on the current survey to your responses on Assessment Day. Once responses have been matched, your student ID will be eliminated from the data file.

Time Required

Participation in this study will require less than 5 minutes of your time.

Risks

The investigators do not perceive more than minimal risks from your involvement in this study, and all information will remain confidential.

Benefits

The objective of this study is to examine how students' sense of identity changes over time throughout their college experience. Because sense of identity has been shown to be related to several desirable behavioral and academic outcomes, knowledge about the development of this construct will help to inform researchers about how the college experience influences identity development. It will also benefit participants in that it will provide an opportunity for each participant to consider their own sense of identity and how it may have changed over their time at JMU.

Confidentiality

The results of this research will be presented at regional and national conferences and in research publications. While individual responses are matched through the use of student ID and recorded online through *Qualtrics* (a secure online survey tool), data is kept in the

strictest confidence. The results of this project will be coded in such a way that the respondent's identity will not be attached to the final form of this study. Aggregate data will be presented representing averages or generalizations about the responses as a whole. All data will be stored in a secure location accessible only to the researchers. Final aggregate results will be made available to participants upon request.

Participation & Withdrawal

Your participation is completely voluntary. You are free to choose not to participate. Should you choose to participate, you can withdraw at any time without consequences of any kind. Should you choose to participate, you may also leave unanswered any items that you would prefer *not* to answer.

Questions about the Study

If you have questions or concerns during the time of your participation in this study, or after its completion or you would like to receive a copy of the final aggregate results of this study, please contact:

Dr. Dena A. Pastor
Associate Professor
Department of Graduate Psychology
James Madison University
pastorda@jmu.edu
(540) 568-1670

Kelli Samonte
2nd Year Master's Student
Psychological Science Program
Department of Graduate Psychology
James Madison University
samontkm@dukes.jmu.edu

Questions about Your Rights as a Research Subject

Dr. David Cockley
Chair, Institutional Review Board
James Madison University
(540) 568-2834
cocklede@jmu.edu
Giving of Consent

I have read this consent form and I understand what is being requested of me as a participant in this study. I freely consent to participate. The investigator provided me with a copy of this form through email. I certify that I am at least 18 years of age. By clicking on the link below, and completing and submitting this confidential online survey, I am consenting to participate in this research.

[insert hyperlink here]

Dr. Dena A. Pastor
Kelli Samonte

Appendix E

The following set of questions deals with how you feel about yourself. Please select the option that indicates the extent to which you agree or disagree with that statement. Please take your time and answer thoughtfully.

	Strongly Disagree	Disagree	Neutral/Undecided	Agree	Strongly Agree
I have a definite sense of purpose in life.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a firm sense of who I am.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a set of basic beliefs and values that guide my actions and decisions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I know what I want out of life.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a clear set of personal values or moral standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't know where I fit in the world.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have specific personal goals for the future.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a clear sense of who I want to be when I am an adult.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please enter your JMU email address (e.g., smithkr@dukes.jmu.edu) so that we may link your responses from those you provided on Assessment Day.

Survey Powered By [Qualtrics](#)

References

- Adams, G.R. , Ryan, J.H., Hoffman, J.J., Dobson, W.R., & Neilsen, E.C. (1985).Ego identity status, conformity behavior, and personality in late adolescence. *Journal of Personality and Social Psychology*, 47, 1091-1104.
- American Association for Higher Education, American College Personnel Association & National Association of Student Personnel Administrators (1998). *Powerful partnerships: A shared responsibility for learning*. Washington, D.C.: Authors
- Archer, S. L. (1982) The lower age boundaries of identity development. *Child Development*, 53, 1551-1556.
- Balistreri, E., Busch-Rossnagel, N. A., & Geisinger, K. F. (1995). Development and preliminary validation of the Ego Identity Process Questionnaire. *Journal of Adolescence*, 18, 179–190.
- Baxter Magolda, M. B. (2003). Identity and learning: Student Affairs' role in transforming higher education. *Journal of College Student Development*, 44 (2), 231-247.
- Bennion, L. D., & Adams, G. R. (1986). A revision of the extended version of the Objective Measure of Ego Identity Status: An identity instrument for use with late adolescents. *Journal of Adolescent Research*, 1, 183–198.
- Chickering, A.W. (1999). Personal qualities and human development in higher education: Assessment in the service of educational goals. In S.J. Messick (ed.), *Assessment in Higher Education* (pp. 13–33). Mahwah, NJ: Lawrence Erlbaum Associates.
- Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: The Guilford Press.
- Erikson, E. H. (1950). *Childhood and society*. New York: Norton.

- Erikson, E. H. (1959). Late adolescence. In D.H. Funkenstein (Ed.), *The student and mental health*, Cambridge, MA: Riverside Press.
- Erikson, E.H. (1968). *Identity: Youth and Crisis*. New York: Norton.
- Erikson, E. H. (1980). *Identity and the life cycle: A reissue*. New York: Norton.
- Faye, C., & Sharpe, D. (2008). Academic motivation in university: The role of basic psychological needs and identity formation. *Canadian Journal of Behavioural Science*, 40, 189-199.
- Field, A. (2009). *Discovering Statistics Using SPSS*. Washington, DC: Sage.
- Grotevant, H. D. (1987). Toward a process model of identity formation. *Journal of Adolescent Research*, 2, 203–222.
- Hoffman, L. (in preparation). *Longitudinal analysis: Modeling within-person fluctuation and change*. Textbook in preparation and under contract with Taylor and Francis
- Jones, R.M. & Hartmann, B.R. (1988). Ego Identity: Developmental differences and experimental substance use among adolescents. *Journal of Adolescence*, 11, 347-360.
- Jones, S. R., & McEwen, M. K. (2000). A conceptual model of multiple dimensions of identity. *Journal of College Student Development*, 41, 405-414.
- Klemp, G. O. (1977). *Three factors of success*. In D. W. Vermilye (Ed.), *Relating Work and Education: Current Issues in Higher Education*. San Francisco: Jossey-Bass.
- Lounsbury, J. W., Saudargas, R. A., & Gibson, L. W. (2004). An investigation of personality traits in relation to intention to withdraw from college. *Journal of College Student Development*, 45, 517-534.

- Lounsbury, J. W., Huffstetler, B. C., Leong, F. T., & Gibson, L. W. (2005). Sense of identity and collegiate academic achievement. *Journal of College Student Development, 46*, 501-514.
- Lounsbury J. W., Saudargas, R. A., Gibson, L. W., & Leong, F. T. (2005). An investigation of broad and narrow personality traits in relation to general and domain-specific life satisfaction of college students. *Research in Higher Education, 46*, 707-729.
- Lounsbury, J. W., & Gibson, L. W. (2011). *Adult Personal Style Inventory*. Knoxville, TN: Resource Associates.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19–40.
- Mallory, M. E. (1989). Q-sort definition of ego identity status. *Journal of Youth and Adolescence, 18*, 399–411.
- Marcia, J. E. (1966) Development and validation of ego identity statuses. *Journal of Personality and Social Psychology, 3*, 551-558.
- Marcia, J. E. (1980). Identity in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 159-187). New York: Wiley.
- Meeus, W., Iedema, J., Helsen, M., & Vollebergh, W. (1999). Patterns of adolescent identity development: Review of literature and longitudinal analysis. *Developmental Review, 19*, 419 – 461.
- Meeus, W. (2011). The study of adolescent identity formation 2000-2010: A review of longitudinal research. *Journal of Research on Adolescence, 21*, 75-94.

- Rosenthal, D. A., Gurney, R. M., & Moore, S. M. (1981). From trust to intimacy: A new inventory for examining Erikson's stages of psychosocial development. *Journal of Youth and Adolescence, 10*, 525–537.
- SAS Institute Inc. (1992). SAS Technical Report P-229, *SAS/STAT Software: Changes and Enhancements, Release 6.07*, Cary, NC: SAS Institute Inc.
- Samonte, K. M., & Pastor D. A. (2011). *Continuous and Categorical Factor Analyses of the Sense of Identity Scale*. Presented at the 42nd annual meeting of the Northeastern Educational Research Association, Rocky Hill, CT.
- Schafer J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods 7*, 147-177.
- Schwartz, S.J. (2001). The evolution of Eriksonian and neo-Eriksonian identity theory and research: A review and integration. *Identity, 1*, 7–58.
- Schwartz, S.J. (2005). A new identity for identity research: Recommendations for expanding and refocusing the identity literature. *Journal of Adolescent Research, 20*, 293–308.
- Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press.
- Toder, N., & Marcia, J. E. (1973). Ego-identity status and response to conformity pressure in college women. *Journal of Personality and Social Psychology, 26*, 287–294.
- Waterman, A. (1982). Identity development from adolescence to adulthood: An extension of theory and a review of research. *Developmental Psychology, 18*, 341-358.

Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods, 14*, 183-201.

Wolfinger, R.D. and Chang, M. (1995). Comparing the SAS GLM and MIXED procedures for repeated measures. *Proceedings of the Twentieth Annual SAS Users Group Conference*, SAS Institute Inc., Cary, NC.