University of Connecticut
OpenCommons@UConn

Master's Theses
University of Connecticut Graduate School

5-15-2020

What is the Impact of Time Lag in Developmental Research? A Re-Analysis of Meta-Analyses Using Lag as Moderator

Rachel Taylor
rachel.m.taylor@uconn.edu

Follow this and additional works at: https://opencommons.uconn.edu/gs_theses

Recommended Citation

Taylor, Rachel, "What is the Impact of Time Lag in Developmental Research? A Re-Analysis of Meta-Analyses Using Lag as Moderator" (2020). *Master's Theses*. 1512.
https://opencommons.uconn.edu/gs_theses/1512

This work is brought to you for free and open access by the University of Connecticut Graduate School at OpenCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of OpenCommons@UConn. For more information, please contact opencommons@uconn.edu.

What is the Impact of Time Lag in Developmental Research?

A Re-Analysis of Meta-Analyses Using Lag as Moderator

Rachel M. Taylor

B.A., Eastern Michigan University, 2018

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

At the

University of Connecticut

2020

Copyright by

Rachel M. Taylor

2020

APPROVAL PAGE

Master of Arts Thesis

What is the Impact of Time Lag in Developmental Research?

A Re-Analysis of Meta-Analyses Using Lag as Moderator

Presented by

Rachel M. Taylor, B.A.

Major Advisor ______________________________
Noel A. Card

Associate Advisor ______________________________
Ryan J. Watson

Associate Advisor ______________________________
Kari L. Adamsons

University of Connecticut
2020

What is the Impact of Time Lag in Developmental Research? A Re-Analysis of Meta-Analyses Using Lag as Moderator

Longitudinal data are essential for evaluating change in developmental science. A survey by Card and Little (2007) reported that 41% of the research published in premier developmental journals is longitudinal. In longitudinal studies, measurements are taken of the same set of individuals at two or more time points. This design allows for the quantification of the strength of two important parameters: Stability and cross-lag prediction. Stability refers to the association between initial and later-time levels of a variable. Its examination helps researchers to understand whether a phenomenon is occurring slowly or rapidly over any given time frame. This, in turn, allows inferences to be drawn about the relative importance of various developmental periods in relation to the construct under examination. For example, studies indicate that early adolescence is a sensitive period for the development of internalizing problems such as depression (Pine et al., 1998). Therefore, one would expect that stability coefficients associated with depression will be lower during this period. Cross-lag prediction involves using initial levels of one variable to predict later levels of another variable. For example, a researcher might utilize cross-lag prediction to examine the strength of the relationship between initial levels of parental support and later-time academic achievement. These associations represent one of the best available methods for approximating causality, where experimentally manipulating variables is impossible or unethical (Little et al., 2009). Conducting sound longitudinal research is, therefore, essential to the continued advancement of developmental science.

Time Lags in Longitudinal Data

Despite the evident importance of conducting methodologically rigorous longitudinal research, little attention has been paid to the time span between longitudinal measurement occasions (time lag). Multiple researchers have highlighted the importance of choosing a theoretically appropriate time lag when designing a longitudinal study. Collins (2006) and Little (2013) both assert that the selection of time lag should be based on (a) the theoretically anticipated rate of change, and (b) the functional form the change process seems likely to take. For example, a researcher might suggest that the ability to speak a language develops primarily during a particular developmental period. In order to appropriately examine this question, measurements must be taken during that developmental period. The number and frequency of measurement occasions should match (a) the speed at which the researcher expects the change process to occur and (b) be sufficient in number to examine the hypothesized functional form of change.

Impact of Lag Length

Previous studies indicate that the length of the time lags found in longitudinal studies may significantly impact the magnitude of the stability coefficients found in that research. Little (2013) illustrated that increased decay in stability coefficients is expected over longer time lags where the change process follows a simplex structure. Specifically, stability can be expected to follow a decreasing pattern equivalent to $r_t = (r_1)^t$, where r_t represents predicted stability over *t* units of time, r_1 indicates stability over one unit of time and t is equivalent to the time lag under examination. For example, where stability is found to be .50 over a one-year time lag, stability over two years will be equivalent to $r_2 = (.50)^2$ = .25. Stability over three years will then be $r_3 = (.50)^3$ = .13, and so forth. Simplex processes bear two key assumptions: (a) that contextual influences are minimal, and (b) that the rate of change under examination is steady across time (Little, 2013). Therefore, the simplex structure might be expected to apply more to some developmental constructs and periods than to others.

The stability of a phenomenon across time may also influence associated models of cross-lag prediction. Specifically, time lag is likely to significantly impact the magnitude and even the direction of the cross-lag coefficients found in longitudinal research. Cole and Maxwell (2003) demonstrated that where levels of a variable are found to be highly stable across time, coefficients of cross-lag prediction may be overestimated if the lag under consideration is longer than the time over which the effect actually occurs. Conversely, where levels of a variable demonstrate lower levels of stability, cross-lag coefficients may be underestimated. Therefore, the expected decrease in stability coefficients demonstrated above may also have a profound influence on cross-lag prediction. Pelz and Lew (1970) ran a series of simulations designed to assess the impact of lag length on coefficients of both stability and longitudinal prediction. Two critical findings came out of this study. First, their results indicated that as lag length increases, coefficients of cross-lag prediction will decrease, approaching zero at a rate dependent upon the strength of associated stability coefficients. Specifically, where stability could be classified as strong, corresponding cross-lag coefficients remained high for a more extended period of time, exceeding the causal interval and leading to over-estimation. Secondly, they found that selecting a lag length that exceeds the true causal interval for a phenomenon can influence the direction of the cross-lag coefficients found in longitudinal research. This is particularly true in the case of negative feedback loops, which are commonly examined in developmental literature.

Methods of Accounting for Time Lag in Longitudinal Research

Multiple methods have been proposed to account for time lag in longitudinal studies. Gollob and Reichardt (1987) advocated including multiple different lag lengths within a study in order to gain a full understanding of a variable's predictive effects. This method of accounting for time lag, which is commonly called variable lag design, involves deliberately including multiple different lag lengths within the same longitudinal study. For example, a researcher using a variable lag design might follow up with a subset of participants at 5 months, another subset at 8 months, and yet another at 10 months. Utilizing such designs allows researchers to examine the uncertain impact of lag length on both stability and cross-lag coefficients. McArdle and Woodcock (1997) demonstrated this variable lag approach and suggested ways it could be implemented practically. The use of variable lag lengths within a study shows promise. However, researchers do not always have the resources or ability to implement variable-lag designs. A longitudinal study may not have the funding necessary to collect data at the desired time or at a theoretically ideal frequency. Alternatively, there may be logistical reasons why collecting data at the desired rate is impossible. For example, a researcher may only be permitted by school administrators to collect data once per year. Finally, it is important to note that many developmental theories do not specify the period over which they expect change to occur. This absence of information means that researchers may be forced to select a lag length blindly. For all of these reasons, the use of convenience lag lengths (operationalized as lag lengths of 6, 12, 24, 36, or 48 months) has become common in longitudinal research. Overuse of such lag lengths may have influenced the results found in primary longitudinal studies.

Lag as Moderator

Although most primary longitudinal studies report a specific, standard, lag length, in practice participants within a study often complete measures at varying times. Selig, Preacher, and Little (2012) proposed the lag as moderator (LAM) approach. This technique capitalizes on existing within-study variability in order to model the impact of different lag lengths on longitudinal effect sizes. Selig and colleagues also suggested that the impact of lag might follow a quadratic rather than a linear functional form. Given that developmental processes take a certain amount of time to unfold, it is likely that an ideal time period exists over which to measure any given phenomenon (Gollob & Reichardt, 1987). Quadratic functional forms show a peak at some point in time, which should correspond with the optimal lag length for measuring any given association. The higher effect sizes associated with this peak indicate the time at with the relationship between the antecedent and the consequence is strongest, which should also be the time period over which a phenomenon is actually occurring.

Selig et al. (2012) demonstrated their approach using data from the Early Head Start Research and Evaluation Study. Existing variation in lag length was already recorded by this study, with the actual time lag between measurements ranging from 5.9 to 17.0 months. In their example, Selig and colleagues assessed lag as a potential moderator for the interaction between home environment and later mental health using both linear and curvilinear functional forms. Their results found significance for both linear and exponential models. However, two potential drawbacks exist for this approach: (a) The range in time lag in a primary study may not be as wide as necessary to detect the optimal time lag for a given process, and (b) it relies on research recording the necessary data to conduct these analyses. Neither of these conditions will likely be met consistently, which potentially limits what LAM can be used to assess.

Lag as Moderator Meta-Analysis

Card (2019) introduced Lag as Moderator Meta-Analysis (LAMMA), a statistical tool that uses between-study variability in lag to assess the impact of different lag lengths on longitudinal effect sizes. LAMMA has several advantages over primary longitudinal studies: (a) It allows for increased variability in lag length, (b) as with most meta-analyses, it includes larger sample sizes and increased statistical power, and (c) it consists of a more heterogeneous overall sample, leading to greater potential for generalizability. It is also not constrained by the same financial and logistical considerations as primary longitudinal studies. Given these considerations, using LAMMA provides the unique opportunity to examine the impact of lag without either a simulation study or collecting new, primary data at variable lag lengths. In cases where theory is nonspecific about when a phenomenon should occur, this gives researchers the opportunity to examine the optimal time lag for measuring any process of interest.

Present Study

Given the evident importance of time lag, it is imperative that we understand how lag length impacts phenomena in developmental science. The present study, therefore, has two aims. First, it utilizes LAMMA to re-analyze the collective data from multiple, previously published, meta-analyses in order to assess the impact of lag length on existing research in developmental science. Utilizing data examining multiple different phenomena has several benefits. It allows us to quantify the potential impact of lag length on our field. It also enables conclusions to be drawn from this research that apply to developmental science generally, rather than to one specific topic. The meta-analyses presented in this thesis contain data from multiple key developmental research areas, including internalizing problems, externalizing problems, personality development, family relationships, peer relationships, adjustment-related topics (e.g., resilience or well-being), adolescent sexual development, and academic achievement.

A second goal of the present study is to assess whether investigators studying a specific phenomenon can use LAMMA to identify when their process of interest may be occurring. As noted earlier, many developmental theories are silent on when they expect their processes of interest to occur. Barring that, there may be sensitive periods for the development of certain phenomena that have not yet been noted in research. If such periods can be identified using LAMMA, researchers can more precisely account for lag in primary longitudinal studies either through variable-lag designs or the LAM approach.

Method

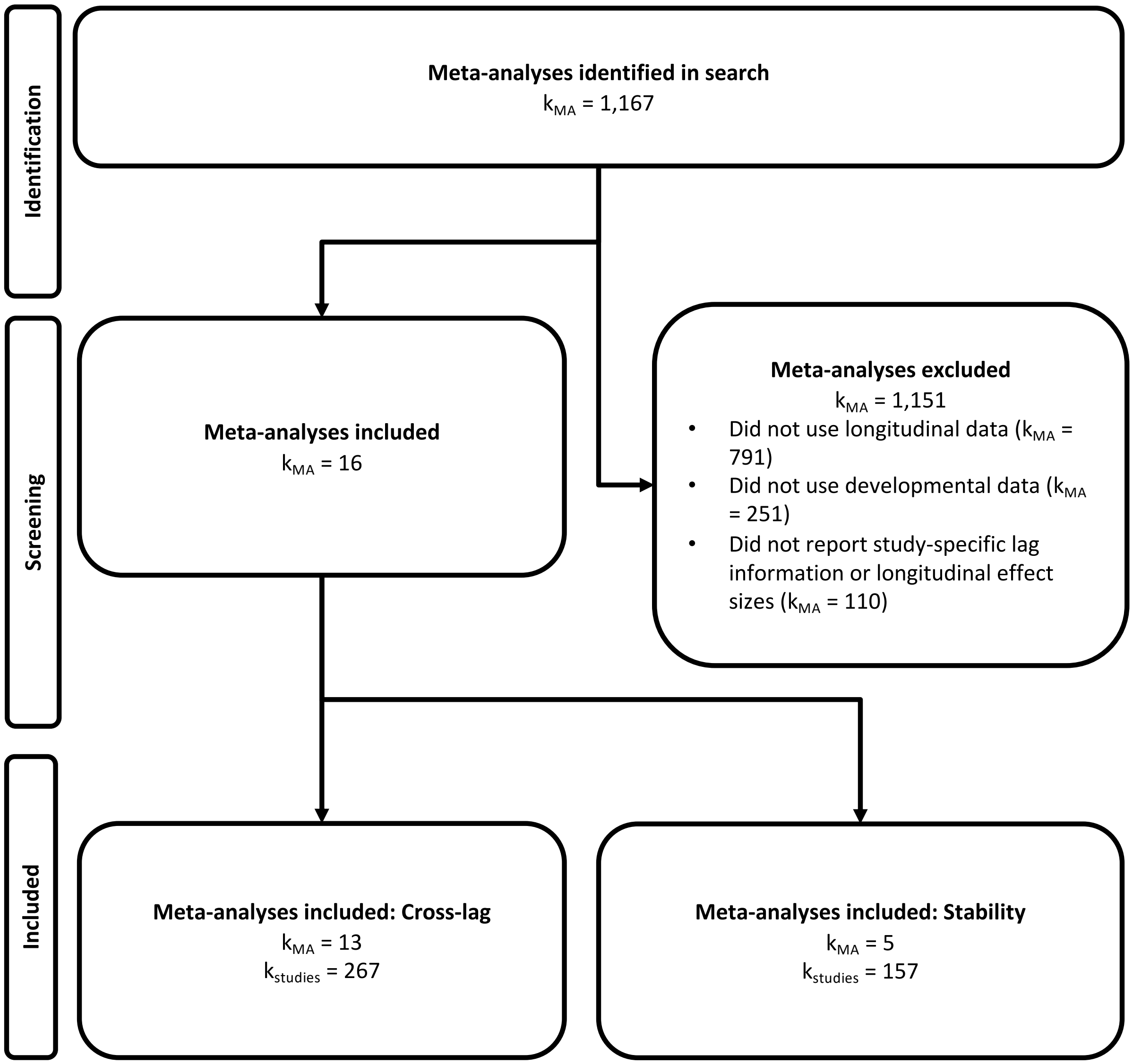
Literature Search and Selection of Studies

Meta-analyses were obtained using two separate search strategies. First, extensive searches of premier developmental journals (including *Child Development*, *Developmental Psychology*, the *International Journal of Behavioral Development*, the *Journal of Research on Adolescence*, and the *Merrill-Palmer Quarterly*) were conducted using the Boolean phrase (“longitudinal study” OR “prospective study”) AND (“meta-analysis” OR “quantitative review”). Searching these specific journals allowed for later comparison of findings on the prevalence of longitudinal studies from Card and Little (2007). Following this initial strategy, PsycINFO was searched using the same Boolean phrase as above. This second strategy ensured that any meta-analyses of longitudinal data that were not found during the initial search were included in this review. The first search was conducted in October 2018, and the second was conducted in February 2019.

Studies were deemed eligible for inclusion if they (a) were meta-analyses of longitudinal data, (b) included study-specific effect sizes (stability, cross-lag coefficients, or both) and lag information, (c) involved a normative (i.e., non-clinical or incarcerated) population, and (d) utilized prospective (as opposed to retrospective) longitudinal data. Effect sizes were calculated as a correlation between Time 1 and Time 2 (Pearson's r). Additionally, lag between studies needed to be reported as chronological time (as opposed to some other unit, e.g., number of interventions). Using this strategy, a total of 16 meta-analyses were deemed eligible for inclusion. Three of these were found in the first literature search, and the remaining 10 were found using the subsequent search of PsycINFO. For a full list of the meta-analyses included in this thesis, see Appendix A. The individual studies included these meta-analyses were analyzed in this study, and are shown in Appendix B.

Figure 1

PRISMA Diagram of Literature Search Process



Note. k_{MA} indicates the number of meta-analyses included. $k_{studies}$ indicates the number of studies coded from those meta-analyses.

Coding of Studies

Both meta-analysis and study-level characteristics were coded. Meta-analysis-level data included descriptive information such as author, title, year published, and journal title. Such information was necessary for assessing (a) the frequency at which meta-analyses of longitudinal data are conducted, and (b) where such studies are most often published. On the study level, sample size was coded and later used for differentially weighting studies. Effect size and type (stability or cross-lag) was also coded, as well as whether researchers controlled for initial levels when calculating cross-lag prediction. However, the latter information was not reported consistently. Lag length (given in months) and average participant age were also coded. Two variables (percent female and percent white) were included in the initial coding scheme but discarded due to insufficient reporting. I coded the entire body of studies twice, first in June 2019 and second in January 2020. There was high intra-coder reliability (97% agreement) between the two waves of coding (see Card, 2012 for full details).

Data Analytic Strategy

Calculating Effect Sizes and Data Preparation

The following steps were taken in analyzing both stability and cross-lag coefficients. One meta-analysis that met inclusion criteria reported some effect sizes as odds ratios. These were converted into standardized mean differences, then into Pearson's r using equations found in Card (2012) and Fleiss and Berlin (2009). Where an individual study was included in multiple meta-analyses, effect sizes were averaged across reports. This approach was methodologically appropriate, given that (a) the meta-analyses were investigating the same or similar constructs, and (b) averaging is generally recommended where multiple effect sizes from a single study qualify for inclusion (see Card, 2012 for further details).

Lag Length

Following coding, variables were checked for skew. Two instances of positive skew were found, specifically in lag length for both the stability and cross-lag samples. Both log and square root transformations have been cited in the literature as correcting well or positive skew because these transformations compress the right side of the distribution towards the left (Osborne, 2002).

Table 1

Lag Lengths

	M (SD)	Range	Skew (SE)

Lag Stability	26.13 (42.27)	1 - 240	3.66
Log Lag Stability	1.18 (.40)	0 - 2.38	.77
Sqrt Stability	4.38 (2.64)	1 - 15.49	2.57
Lag Cross-lag	28.24 (51.18)	.36 - 456	4.75
Log Lag Cross-lag	1.09 (.56)	-.44 - 2.66	0.02
Sqrt Cross-lag	4.32 (3.09)	.60 - 21.35	2.20

Note. Log = log transformation used. Sqrt = Square root transformation used. The unit for lag is given in months.

The log transformation successfully corrected for positive skew. Therefore all analyses were conducted with log-transformed lag (see Appendix C for a graphical representation of the log lag transformation).

LAMMA

The following analyses were conducted for both the stability and cross-lagged samples. Each effect size was transformed into Fisher's *Zr* prior to analysis. Two potential models, linear and quadratic, were examined. In the linear models, effect sizes were regressed onto lag, weighted by traditional inverse variance ($W = N - 3$). Prior to estimating the quadratic model, lag length was centered using the methods found in Card (2019). Centering lag reduced nonessential collinearity (Aiken & West, 1991). The centered lag variable was then squared to create a quadratic term. Both the centered lag and quadratic lag variables were then used as predictors of effect sizes, weighted by inverse variance. Heterogeneity among effect sizes was evaluated using Q, a statistic that describes the variability accounted for by each model. A fixed effects model was used to conduct analyses in both Stata and SPSS.

Results

A total of 1,167 potential meta-analyses were reviewed for inclusion. Of these, 791 were not meta-analyses of longitudinal data. A range of studies fell into this category, including primary studies, simulation studies, and theoretical papers. A number of meta-analyses (k = 251) were also excluded because they did not utilize developmental data, which was operationalized as focusing on the development of physical, cognitive, social, intellectual, perceptual, personality, or emotional processes.

Meta-analyses examining medical questions, clinical trials, and incarcerated were also excluded from the present study. Finally, 110 studies were excluded because they did not report study-specific lag or longitudinal effect sizes. A final sample of 16 meta-analyses was included in this study. Of these, 5 ($k_{studies}$ = 157, *N* = 45,443) met inclusion criteria for stability while 13 ($k_{studies}$ = 267, *N* = 190,082) met criteria for cross-lag associations (Figure 1).

Average age weighted by sample size was calculated for both the stability and cross-lag samples. The majority of the sample of studies was utilized in these calculations. However, a number of studies did not report the necessary data in both the stability ($k_{studies}$ = 15) and cross-lag ($k_{studies}$ = 5) samples. Therefore, the average age of both populations should be treated as an approximation.

Table 2

Summary of Characteristics of Included Studies

	Stability		Cross-lag	
	M (SD)	Total	M (SD)	Total
Meta-analyses (k_{MA})		5		13
Studies ($k_{studies}$)		157		267
Participants (N)	289.45 (952.05)	45,443	711.92 (1585.05)	190,082
Sample age (years)	7.52 (8.98)		14.04 (9.85)	
Lag length (months)	26.23 (42.27)		28.24 (51.18)	

Note. Average age refers to a participant's average age at Time 1. M = mean. SD = standard deviation.

Frequency of Longitudinal Meta-Analyses

As noted in the Introduction, Card and Little (2007) surveyed five leading developmental journals in order to assess the prevalence of longitudinal studies. The search of the current thesis of these same journals indicated that from 2013 to 2018, they collectively published 85 meta-analyses. The total number of longitudinal meta-analyses found in these journals in this span (*k* = 10) indicates that 11.8% of meta-analyses used longitudinal data.

LAMMA Analyses

Linear Effect of Lag on Stability

Transformed stability (*Zr*) was regressed onto log-transformed lag using a fixed effects model. Results for this analysis indicate that lag is a significant moderator (*Zr* = -.14, $p < .01$). Specifically, the linear model indicated that strength in stability will decrease in accordance with the following equation:

$$\widehat{Zr} = .717 - .14 * loglag$$

In accordance with the above results, this model accounted for a significant amount of heterogeneity in effect sizes ($Q_{(1)} = 100.01$, $p < .001$), with the overall model accounting for 6.26% of the overall heterogeneity in effect sizes.

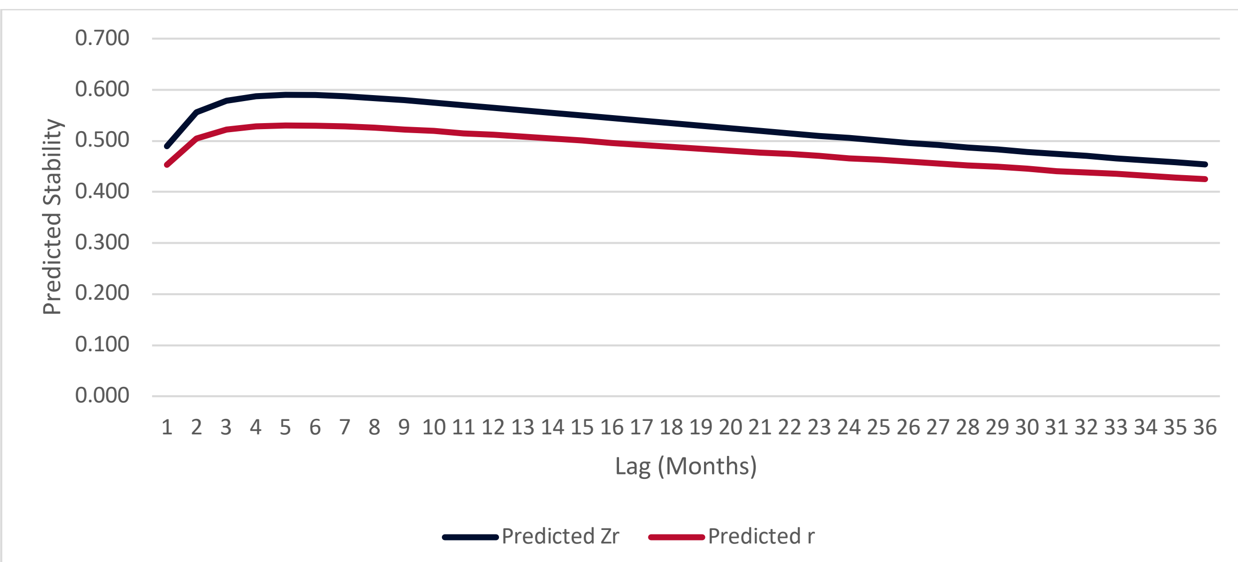
Quadratic Effect of Lag on Stability

The quadratic model for stability also indicated significance for both linear (*Zr* = -.14, $p < .01$) and quadratic terms (*Zr* = -.19, $p < .01$). This yielded the following predictive equation (see Figure 2).

$$\widehat{Zr} = .565 - .14 * centered\ lag - .19 * quadratic\ lag$$

Figure 2

Predicted Quadratic Decay in Stability Using Back-transformed Lag

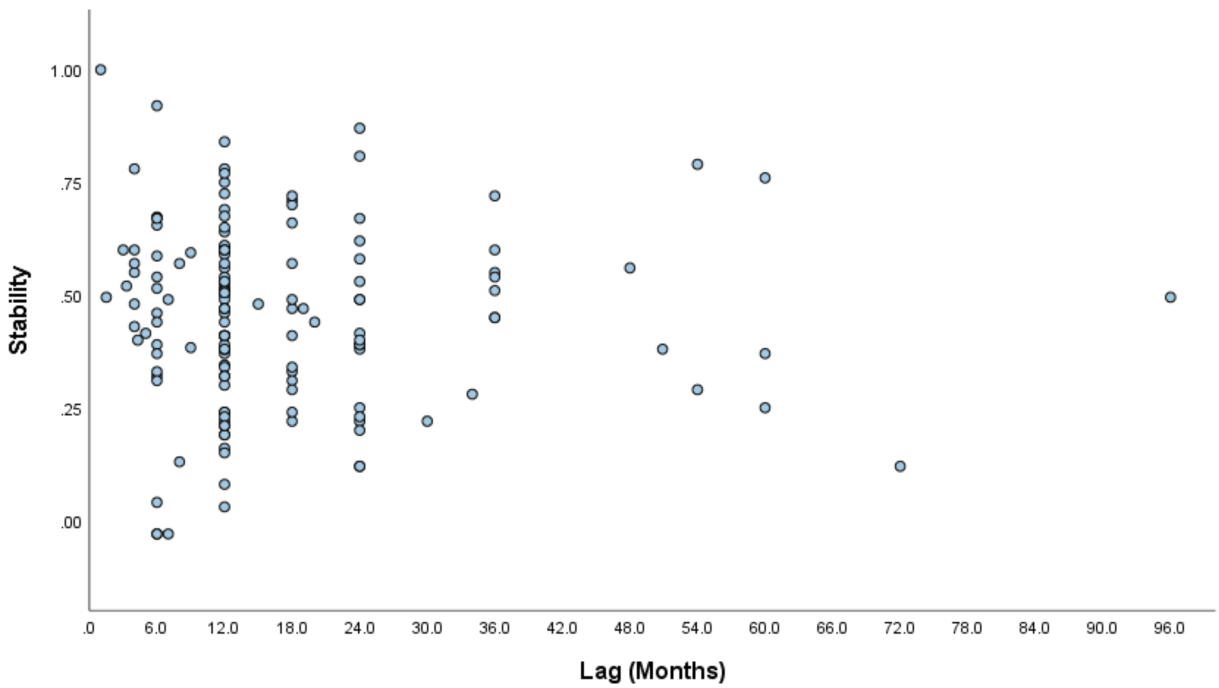


Note. The blue line represents predicted stability (*Zr*) and, the red represents back-transformed Pearson's r. For similar graphs showing predicted stability values (a) prior to back-transforming log lag, and (b) displaying the predicted linear models, see Appendix E.

This model also indicated a significant amount of heterogeneity in effect sizes ($Q_{(2)} = 153.62$, $p < .001$, with the overall model accounting for 9.62% of the overall heterogeneity in effect sizes. This represents a substantial increase in terms of the amount of heterogeneity accounted for when compared to the linear model. The quadratic model for lag, therefore, represents a better fit for the stability data.

Figure 3

Scatterplot of Stability (r) Regressed onto Lag



Note. Data in the above plot is truncated by a number of studies (k = 7) in order to display the preponderance of convenience lag lengths.

The above figure shows stability (*r*) regressed onto lag length (months). As shown, convenience lag lengths were prevalent (see the multitude of time points at 6, 12, 24, 36, and 48 months). In the stability sample, 65.6% lag lengths fell into that category. This may have introduced error into the above analyses by disproportionately weighting certain time points. This may indicate that lag length had a more profound impact on stability coefficients than the above analyses were able to demonstrate.

Linear Effect of Lag on Cross-lag Prediction

A fixed effects model was also used to assess the impact of lag on cross-lag prediction. Results indicated that lag was a significant moderator (Zr = -0.27, p < .05). These analyses indicated the overall predictive equation:

$$\widehat{Zr} = .104 - .027 * loglag$$

In accordance with the above results, this model indicated that a significant amount of heterogeneity in effect sizes ($Q_{(1)}$ = 38.74, p < .05), with the overall model accounting for 1.99% of the overall heterogeneity in effect sizes.

Quadratic Effect of Lag on Cross-lag Prediction

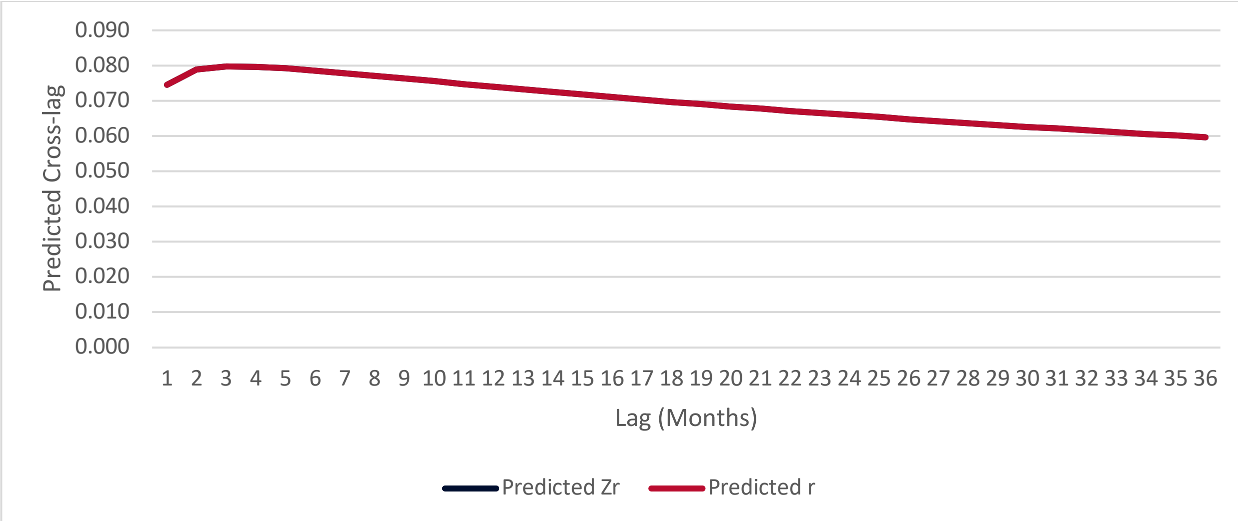
As with stability, the quadratic model also indicated significance for cross-lag prediction for both linear (Zr = -.021, p < .01) and quadratic terms (Zr = -.019, p < .01). This model indicated that strength in cross-lag will decrease in accordance with the following equation:

$$\widehat{Zr} = .074 - .021 * centered\ lag - .019 * quadratic\ lag$$

A graphical depiction of the above equation is illustrated in Figure 4. This model also predicted a significant amount of heterogeneity in effect sizes ($Q_{(2)}$ = 52.06, p < .05), with the overall model accounting for 2.67% of the overall heterogeneity in effect sizes. Figure 3 shows values predicted from this model.

Figure 4

Predicted Quadratic Decay in Cross-lag Using Back-Transformed Lag

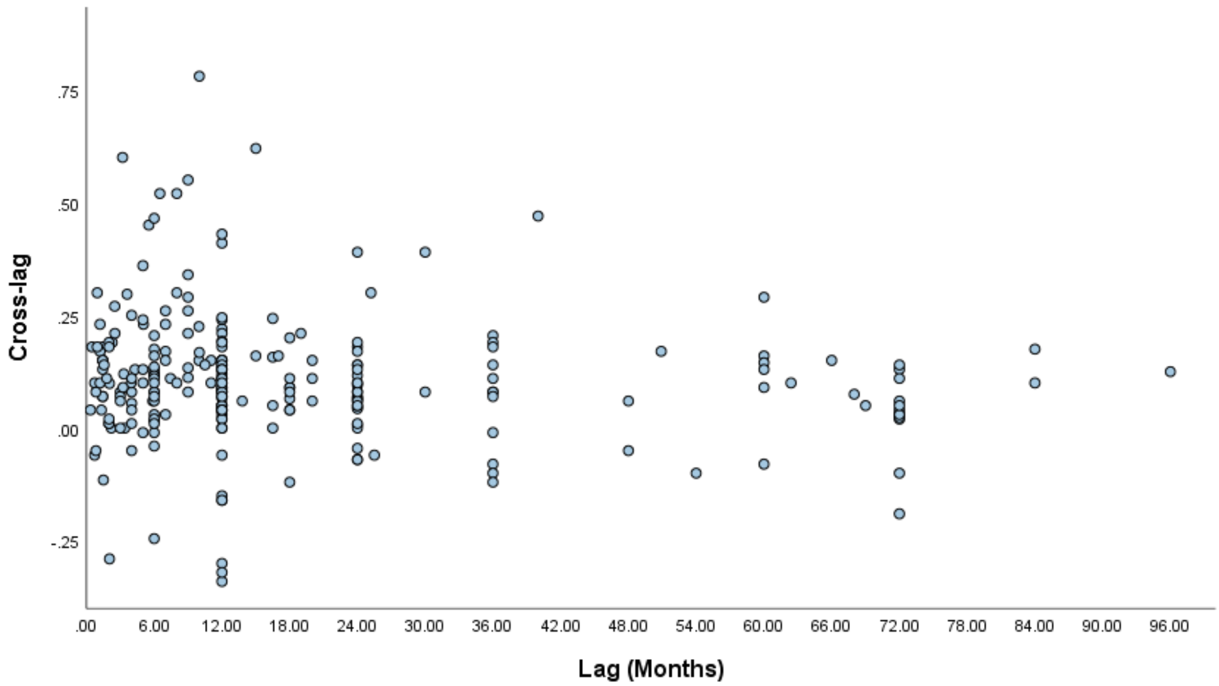


Note. In the above graph, the values for predicted Zr and Predicted r are very similar, so the curve for predicted r obscures the curve for predicted Zr. For graphs showing predicted cross-lag values (a) prior to back-transforming log lag, and (b) displaying the predicted linear models, see Appendix F.

It is notable that, overall, cross-lag coefficients begin lower and decrease at a slower rate with lag than stability coefficients.". However, this is likely because the association between a variable and itself is generally stronger than the association between a variable and another variable. Additionally, the variety of phenomena analyzed in this meta-analysis may have introduced more heterogeneity in effect sizes for cross-lag prediction than for stability. Specifically, the variety of topics analyzed increased the likelihood that variables that do not predict each other (nonsignificant results) were included, drawing down overall results.

Figure 5

Scatterplot of Cross-lag (r) Regressed onto Lag



Note. Data in the above plot is truncated by a number of studies (k = 13) in order to display the preponderance of convenience lag lengths.

Convenience lag lengths were less prevalent within the cross-lag sample than the stability sample. Specifically, convenience lag lengths accounted for 32.9% of all lag lengths for cross-lag prediction. This almost certainly impacted results, likely leading the impact of lag to be under-estimated in this study.

Discussion

The present meta-analysis investigated the impact of lag length on both stability and cross-lag associations in longitudinal developmental research. It utilized a broad range of previously existing meta-analyses, which allows these results to be generalized across developmental science. The results presented in this thesis may help to guide (a) how lag is accounted for in future meta-analyses of longitudinal data, and (b) how future primary longitudinal studies should go about choosing a lag length.

LAMMA Models

Analyses indicated that both the linear and quadratic models for stability were significant and that increased lag length will likely be associated with lower stability coefficients. These results indicate that both linear and quadratic functions may adequately model decline in stability. The quadratic model for stability indicated a peak in the strength somewhere between four and six months. However, this model illustrates a general trend only and should not be used to justify assessing stability over time lags within that range. Rather, it should serve as a benchmark indicating what meta-analysts measuring specific phenomena might want to assess in future studies. While the observed decrease in stability does not follow the expected decrease for a simplex structure, this result is to be expected. Given that data from the included meta-analyses were conducted on a variety of different constructs during variable developmental periods, it does not make sense to assume that the rates of change being considered would be stable across time.

Similar results were found for the cross-lag prediction models, where both functional forms were also found to be significant. This model indicated a potential peak in the strength of cross-lag prediction ranging approximately from two to four months. It is noteworthy that this value is much shorter than the majority of lag lengths used in longitudinal studies. However, this result should not be taken as an indication that future studies should utilize time lags within that range. Developmental processes occur at varying rates, and at varying times. As such, there can be no ideal lag length for the field as a whole. Taken together, the results for both stability and cross-lag support the idea that future researchers who use LAMMA to model the impact of lag should investigate multiple functional forms. Indeed, the quadratic models for both stability and cross-lag prediction were found to predict an increased percentage of heterogeneity when compared to linear models of the same effect sizes.

The primary studies used in this meta-analysis made extensive use of convenience lag lengths in both the stability and the cross-lag samples. Although it is unclear what impact this may have had on the above results, it seems likely that they biased findings in some way, either through (a) the collective impact of primary studies' inappropriate choice of lag length, or (b) simply by giving more weight to certain time points than to others. This is particularly true for the stability sample, where convenience lag lengths accounted for the majority (66.5%) of the lag lengths included in analyses. The prevalence of such lag lengths is concerning, as it may have biased many primary longitudinal studies.

Prevalence of Longitudinal Meta-Analyses

Longitudinal meta-analyses make up a small proportion of the meta-analyses published each year. However, meta-analyses of longitudinal data are obviously valuable in that they enable researchers to draw associations between variables across time with increased sample size, power, and sample diversity. While the results of the present study do not support definitive conclusions about what lags to use to measure particular developmental phenomenon, its findings do indicate that LAMMA would be helpful to design the lags to use in future primary studies. Specifically, if future meta-analyses focusing on specific phenomena take lag into account, then that information can inform future primary longitudinal studies how and when to space measurement occasions. This will lead to an overall improvement in the quality of results found in longitudinal research, and so in developmental science as a whole.

Limitations

A major strength of the present study was that it included a wide array of developmental constructs, thereby ensuring that these results can be applied to developmental research as a whole. However, this decision may also have introduced additional heterogeneity, thereby decreasing the impact of lag that this study was able to show. It is therefore likely that future meta-analyses of longitudinal data that analyze a single construct will find results indicating that the impact of lag length was greater than it appeared in the present study. There are also potential limits to the generalizability of these results, particularly in the case of stability. Although the present study analyzed data from a large number of primary studies, they were coded from a relatively small sample of meta-analyses. This may be particularly true of the stability sample, which included five meta-analyses overall. While this number is unlikely to have impacted results quantitatively (i.e., the number of included studies was more than sufficient to run moderation analyses), this does mean that a smaller number of developmental constructs were included in that sample relative to the cross-lag analyses. This may limit the generalizability of those results. Finally, it is critical to emphasize that the results presented in this thesis are meant to inform how future meta-analyses of longitudinal data account for lag: These results should not be used to guide what lag length is used to measure a specific phenomenon.

The present study utilized a fixed-effects model, which is justifiable given that it exclusively examined a normative population. However, it may be that our sample was more heterogeneous than the included meta-analyses indicated. Therefore, it may be that a random-effects model would be a better fit for these data. Results were additionally limited by the data provided in the original meta-analyses. For example, it was impossible to correct for measurement effort given that none of the included meta-analyses reported study-specific values of Cronbach's alpha. Very few studies ($k = 2$) also controlled for initial levels of variables, meaning that error may have been introduced because studies did not the impact variables had on themselves into account (see Little, 2013, for further information on controlling for initial levels of a variable).

Future Directions

Future analyses should examine the impact of age in conjunction with the impact of lag length. Developmental periods are widely theorized to be enormously important with certain age ranges (e.g., childhood, adolescence) being more sensitive for the development of certain phenomena than others. Including average participant age in models will allow researchers to (a) take developmental period into account when examining lag, and (b) allow researchers to compare the relative importance of various periods for the development of any given construct. These data can then be compared to (a) other developmental periods, or (b) the general population. This should allow researchers a medium with which to establish the relative importance of various developmental periods.

Conclusion

In sum, results from the present meta-analysis demonstrate the necessity of accounting for time lag both in meta-analyses of longitudinal data and in primary longitudinal studies. Longitudinal meta-analyses have the potential to provide valuable information about both the functional form, the impact of lag takes, and the optimal time period over which to measure any given phenomenon. Using these results to inform how we select time lags in primary longitudinal studies will decrease lag-related error, thereby improving the overall validity of the results found in longitudinal studies and in developmental science as a whole.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA: SAGE Publications.

*Braga, T., Goncalves, L.C., Basto-Pereira, M., & Maia, A. (2017). Unraveling the link between maltreatment and juvenile antisocial behavior: A meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior, 33,* 37-50. http://dx.doi.org/10.1016/j.avb.2017.01.006.

Card, N. A. (2012). *Applied meta-analysis for social science research*. New York, NY: Guilford Press.

Card, N. A. (2019). Lag as moderator meta-analysis : A methodological approach for synthesizing longitudinal data. *International Journal of Behavioral Development, 43*(1), 80–89. https://doi.org/10.1177/0165025418773461.

Card, N. A., & Little, T. D. (2007). Longitudinal modeling of developmental process. *International Journal of Behavioral Development*, *31*(4), 297–302. https://doi.org/ 10.1177/0165025407077750.

Cole D. A., & Maxwell S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology 112*, 58–577. https://doi.org/10.1037/0021-843X.112.4.558.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, *57*, 505–528. https://doi.org/10.1146/annurev.psych.57.102904.190146.

*Colonnesi, C. Geert, J., Stams, J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review 30*(4), 352-366. https://doi.org/10.1016/j.dr.2010.10.001.

Field, A. P. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. Thousand Oaks, CA: SAGE Publications.

*Fraley, C. R. (2002). Attachment Stability From Infancy to Adulthood: Meta-Analysis and Dynamic

Modeling of Developmental Mechanisms. *Personality and Social Psychology Review*, 6(2), 123–151. https://doi.org/10.1207/S15327957PSPR0602_03.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*(1), 80–92.

Gollob, H. F. & Reichardt, C. S. (1991). Interpreting and estimating indirect effects assuming time lags really matter. In: Collins, L.M.; Horn, JL., editors. *Best methods for the analysis of change: Recent advances, unanswered questions, future directions.* Washington D. C: American Psychological Association.

*Huang, C. (2015). Relation between attributional style and subsequent depressive symptoms: A systematic review and meta-analysis of longitudinal studies. *Cognitive Therapy and Research, 39*(6), 721–735. https://doi.org/10.1007/s10608-015-9700-x.

*Koletić, G. (2017). Longitudinal assocations between the use of sexually explicit material and adolescents' attitudes and behaviors: A narrative review of studies. *Journal of Adolescence, 57*, 119- 133. http://dx.doi.org/10.1016/j.adolescence.2017.04.006.

*Kuykendal, L. T., Tay, L., & Ng, V. (2015). Leisure engagement and subjective well-being: A meta-analysis. *Psychological Bulletin 141*(2), 364-403. http://dx.doi.org/10.1037/a0038508.

*La Paro, K. M., & Pianta, R. C. (2000). Predicting Children's Competence in the Early School Years: A Meta-Analytic Review. *Review of Educational Research*, 70(4), 443–484. https://doi.org/10.3102/00346543070004443

*Larzelere, R.E., Gunnoe, M.L., & Ferguson, C.J. (2018). Improving causal inferences in meta-analyses of longituidnal studies: Spanking as an illustration. *Child Development, 89*(6), 2038-2050. https://doi.org/10.1111/cdev.13097.

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford.

Little, T. D., Card, N. A., Preacher, K. J., & McConnell, E. (2009). Modeling longitudinal data from research on adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 15–54). 3rd edition. New York, NY: Wiley.

McArdle, J. J., & Woodcock, R. W. (1997). Expanding test–retest designs to include developmental time-lag components. *Psychological Methods,* 2(4), 403–435. https://doi.org/10.1037/1082-989X.2.4.403.

Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research, and Evaluation, 8*(8), https://doi.org/10.7275/4vng-5608.

Pelz, D. C.. & Lew, R. A. (1970). Heise's causal model applied. In: Borgatta, EF.; Bohrnstedt, G., editors. *Sociological Methodology*. San Francisco, CA: Jossey-Bass;. p. 28-37.

*Pouwels, J. L., Salmivalli, C., Saarento, S., van den Berg, Y. H. M., Lansu, T. A. M., & Cillessen, A. H. N. (2018). Predicting adolescents' bullying participation from developmental trajectories of social status and behavior. *Child Development*, 89(4), 1157–1176. https://doi.org/10.1111/cdev.12794.

*Reijntjes A., Kamphuis, J.H., Prinzie, P., Boelen, P.A., van der Schoot, M., & Telch, M.J. (2011). Prospective linkages between peer victimization and externalizing problems in children: a meta-analysis. *Aggressive Behavior*, 37(3), 215-222. https://doi.org/10.1002/ab.20374.

*Reijntjes, A. & Kamphuis, J., Prinzie, P. & Telch, M. (2010). Peer victimization and internalizing problems in children: A meta-analysis of longitudinal studies. *Child abuse and neglect,* 34, 244-52. https://doi.org/10.1016/j.chiabu.2009.07.009.

*Riglin, L., Petrides, K. V., Frederickson, N., & Rice, F. (2014). The relationship between emotional problems and subsequent school attainment: A meta-analysis. *Journal of Adolescence 37*(4), 335-346. https://doi.org/10.1016/j.adolescence.2014.02.010.

Selig, J. P., Preacher, K. J. & Little, T. D. (2012). Modeling time-dependent assocation in longitudinal data: A lag as moderator appraoch. *Multivariate Behavioral Research, 47*(5): 697-716, https://doi.org/10.1080/00273171.2012.715557.

*Smith, M., Vidovic, V., Sherry, S.B., Stewart, S.H., & Saklofske, D.H. (2018). Are perfectionism dimensions risk factors for anxiety symptoms? A meta-analysis of 11 longitudinal studies. *Anxiety, Stress, and Coping*, *31*(1), 4-20. https://doi.org/10.1080/10615806.2017.1384466.

*Smith, M. M., Sherry, S. B., Rnic, K., Saklofske, D. H., Enns, M., & Gralnick, T. (2016) Are perfectionism dimensions vulnerability factors for depressive symptoms after controlling for neuroticism? A meta-analysis of 10 longitudinal studies. *European Journal of Perssonality*, 30, 201– 212. https://doi.org/10.1002/per.2053.

*Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review, *Educational Psychologist*, 39:2, 111-133. https://doi.org/10.1207/s15326985ep3902_3.

*Wardle, J., Chida, Y., Gibson, E. L., Witaker, L .& Steptoe, A. (2011). Erratum: Stress and Adiposity: A Meta-Analysis of Longitudinal Studies. *Obesity*, 19: 1315-1315. https://doi.org/10.1038/oby.2010.268.

* = Meta-analyses included in the present study

Appendix A

Table 3
Included Meta-analyses

Author	Year	Academic Journal	Stability	Cross-lag
Braga et al.	2017	Aggression and Violent Behavior	No	Yes
Colonnesi	2010	Developmental Review	No	Yes
Fraley	2002	Psychology Review	Yes	No
Huang and Chiungiung	2015	Cognitive Therapy and Research	No	Yes
Koletić and Goran (2017)	2015	Journal or Adolescence	No	Yes
Kuykendall	2015	Psychological Bulletin	No	Yes
La Paro and Pianta	2000	Review of Educational Research	Yes	No
Larzelere et al.	2018	Child Development	Yes	Yes
Pauwles et al.	2016	Developmental Review	Yes	No
Reijntjes et al.	2011	Aggressive Behavior	No	Yes
Reijntjes et al.	2010	Child Abuse and Neglect	No	Yes
Riglin et al.	2014	Journal or Adolescence	No	Yes
Smith et al.	2016	European Journal of Personality	Yes	Yes
Smith et al.	2018	Anxiety, Stress, and Coping: An International Journal	Yes	Yes
Valentine et al.	2004	Educational Psychologist	No	Yes
Wardle et al.	2011	Obesity	No	Yes
Total k			**6**	**13**

Note. Stability and cross-lag columns indicate whether the relevant effect size was reported in an included meta-analysis.

Appendix B

Table 4

Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Braga et al. (2017)					
Bank and Burraston (2001)	181	11	66	-	0.15
Brezina (1998)	1502	15	18	-	0.09
Bright and Johnson-Reid (2008)	1607	9	72	-	0.03
Chapple et al. (2005)	942	4	144	-	0.04
Fagan and Wright (2011)	651	13.5	36	-	0.07
Jovev et al (2013)	664	13.5	36	-	0.14
Kaufman (2015)	205	12.5	25.2	-	0.30
Kazermian et al. (2011)	950	15.2	36	-	0.11
Knutson and DeGarmo (2004)	375	9	72	-	0.02
Knutson and DeGarmo (2004)	361	10.5	60	-	0.16
McCabe et al. (2005)	310	6.5	12	-	0.22
Rebellon and Van Gundy (2005)	397	13.6	24	-	0.05
Reyes et al. (2015)	1113	14	24	-	0.07
Salzinger et al. (2007)	1759	14	12	-	0.05
Shenk et al. (2010)	153	10.5	60	-	0.15
Simmel (2007)	144	11.1	84	-	0.18
Sternberg et al. (2006)	276	4.9	72	-	0.14
Stevens (2012)	83	10.7	62.4	-	0.10
You and Lim (2015)	306	10.8	36	-	0.21
Colonnesi et al. (2010)					
Aureli et al. (2008)	18	1	5	-	0.24
Bates et al. (1979)	25	0.83	2.5	-	0.27
Blake et al. (2003)	12	0.82	2.2	-	0.00
Brooks and Meltzoff (2008)	32	2.67	6.5	-	0.52
Butterworth and Morissette (1996)	27	0.92	3.4	-	0.00
Camaioni et al. (1991)	23	1	8	-	0.52
Carpenter et al. (1998)	24	1	12	-	0.08
Colonnesi et al. (2008)	35	1.13	25.5	-	-0.06
Delgado et al. (2002)	47	1.25	9	-	0.55
Desrochers et al. (1995)	25	1	12	-	0.43
Fasolo and d'Odorico (2002)	44	1.67	10	-	0.78
Markus et al. (2000)	21	1.25	9	-	0.21
Morales et al. (2000)	22	1.25	15	-	0.62
Mundy and Gomes (1998)	24	1.32	3.2	-	0.60
Mundy et al. (2003)	29	1.33	8	-	0.30
Mundy et al. (2007)	72	1.13	10.5	-	0.14
Perucchini and Plescia (2005)	50	0.92	9	-	0.26
Rowe and Goldin-Meadow (2008)	50	1.17	40	-	0.47
Fraley (2002)					
Ammaniti et al. (1996)	20	1	48	0.56	-

Table 4 (continued)
Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Fraley (2002) (continued)					
Belsky et al. (1996)	90	1	6	0.56	-
Belsky et al. (1996)	120	1	6	-0.03	-
Egelund and Farber (1984)	189	1	6	0.04	-
Egelund and Sroufe (1981)	25	1	6	0.32	-
Egelund and Sroufe (1981)	32	1	6	0.33	-
Frodi et al. (1985)	38	1	8	0.67	-
Goossens et al. (1986)	9	1	1	0.13	-
Hamilton (2000)	30	1	216	1.00	-
Howes and Hamilton (1992)	23	1	7	0.50	-
Howes and Hamilton (1992)	72	1	30	0.49	-
Howes and Hamilton (1992)	89	1	36	0.22	-
Jacobsen et al. (1997)	32	1	6	0.45	-
Jacobsen et al. (1997)	32	1	60	0.31	-
Lyons-Ruth et al. (1991)	46	1	6	0.37	-
Main (In Press)	38	1	216	-0.03	-
Main and Weston (1981)	15	1	6	0.50	-
Main et al. (1985)	40	1	60	0.46	-
Owen et al. (1984)	59	1	8	0.76	-
Schneider-Rosen et al. (1985)	29	1	6	0.57	-
Thompson et al. (1982)	43	1	7	0.39	-
Vaughn et al. (1979)	100	1	6	-0.03	-
Wartner et al. (1994)	39	1	54	0.37	-
Waters (1978)	50	1	6	0.79	-
Waters et al. (2000)	50	1	240	0.92	-
Weinfeld et al. (2000)	57	1	216	0.45	-
Zimmerman et al. (1997)	43	1	180	0.10	-
Huang and Chiungjung (2015)					
Abela et al. (2009)	342	14.13	1.44	-	0.18
Abela and Payne (2003)	314	11.14	1.44	-	0.15
Abela and Sarin (2002)	79	12.25	2.28	-	0.19
Alvardo (1988)	109	19.5	1.44	-	0.15
Asdigian (1993)	247	19.5	0.72	-	0.10
Bramlette (1998)	115	19.3	0.48	-	0.18
Brozina and Abela (2006)	418	10.5	1.44	-	0.13
Cole et al. (2011)	100	8.51	30	-	0.08
Colnley et al. (2001)	130	8.2	7.44	-	0.11
Daniels (1999)	38	43.07	3	-	0.07
Edelman et al. (1994)	94	19.5	0.72	-	-0.06
Frantom (1994)	71	21.6	1.44	-	0.07
Gibb and Abela (2008)	105	9.82	12	-	-0.16
Gibb and Abela (2008)	106	12.27	24	-	0.18

Table 4 (continued)
Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Huang and Chiungjung (2015) (continued)					
Gibb et al. (2006)	417	9.77	5.88	-	0.06
Gibb et al. (2012)	100	9.97	6	-	0.03
Grazioli and Terry (2000)	57	28.81	1.44	-	0.07
Guerry (2008)	101	13.51	2.52	-	0.21
Hamilton (1982)	20	36.6	6	-	-0.04
Han (1995)	62	37	0.84	-	0.08
Hankin et al. (2001)	153	16.18	1.2	-	0.17
Hankin et al. (2001)	117	16.18	1.2	-	0.23
Hilsman and Garber (1995)	414	11.39	0.36	-	0.04
Johnson (1992)	100	19.5	0.96	-	0.30
Johnson and Miller (1990)	80	19.5	0.96	-	0.18
Joiner (2000)	34	14.33	2.04	-	-0.29
Kleim et al. (2012)	183	35.14	5.52	-	0.45
Kleiman et al. (2012)	209	20.51	0.84	-	-0.05
Kouros et al. (2013)	240	11.86	72	-	0.04
Kuperman (1991)	63	41	1.44	-	0.15
Lewinsohn et al. (2001)	1507	16.6	13.8	-	0.06
Martin (1986)	305	19.61	2.04	-	0.09
McCarty et al. (2007)	331	12	12	-	0.10
McQuade et al. (2011)	88	9.6	12	-	0.02
Metalsky and Joiner (1992)	152	19.5	1.2	-	0.10
Morris and Tiggeman (1999)	247	22.04	9.96	-	0.23
Nolen-Hoeksema (1986)	168	9	12	-	0.15
O'Donnell et al. (2010)	88	10.74	24	-	0.05
Panak and Garber (1992)	521	9	11.04	-	0.15
Pomerantz (2001)	806	11.69	6	-	0.12
Possel and Thomas (2011)	311	23.27	1.8	-	0.11
Priester and Clum (1992)	269	19.5	3.6	-	0.29
Prinstein and Aikins (2004)	158	16.31	5.04	-	0.23
Quevedo (2008)	170	11	69	-	0.05
Reilly et al. (2012)	140	19.5	1.32	-	0.04
Robinson et al. (1995)	239	12	9.96	-	0.15
Rueger and Malecki (2011)	257	13.2	3.96	-	0.10
Sanjuan and Magallares (2009)	101	37.01	1.56	-	0.14
Southall and Roberts (2002)	115	16.5	3.24	-	0.09
Spence et al. (2002)	733	12.91	12	-	0.13
Stevens and Prinstein (2005)	398	12.7	11.04	-	0.10
Stone et al. (2010)	417	18.14	6	-	0.12
Syzdek and Addis (2010)	62	38	3	-	0.06
Williams (1998)	122	19.8	12	-	0.15

Table 4 (continued)
Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Koletić and Goran (2017)					
Baams et al. (2014)	444	13.9	6	-	0.47
Brown and L'Engle (2009)	967	13.6	24	-	0.12
Doornwaard et al. (2015)	1132	13.9	6	-	0.16
Hennessy et al. (2010)	506	15	12	-	-0.34
van Oosten and Vandenbosch (2015)	1765	15	6	-	0.13
Peter and Valkenburg (2008)	962	16.8	6	-	0.14
Peter and Valkenburg (2011)	1445	14.5	6	-	0.12
Vandenbosch et al. (2013)	639	14.8	6	-	0.08
Ybarra et al. (2011)	1159	12.6	12	-	0.99
Kuykendall (2015)					
Pinquart and Schindler (2009)	669	-	12	-	0.11
Pinquart and Schindler (2009)	938	-	12	-	0.15
Powdthavee (2009)	11350	-	12	-	0.15
Shin and You (2013)	1594	-	12	-	0.24
Shin and You (2013)	1594	-	12	-	0.08
La Paro and Pianta (2000)					
Agostin and Bain (1997)	135	5	12	0.32	-
Badian (1994)	117	4	24	0.38	-
Beringer et al. (1990)	42	5	12	0.53	-
Berninger (1986)	45	5	12	0.41	-
Berninger and Alsdorf (1989)	27	5	12	0.03	-
Berninger et al. (1988)	28	5	12	0.21	-
Campbell et al. (1986)	51	4	24	0.62	-
Campbell et al. (1991)	73	4	12	0.39	-
Catts (1991)	41	5	12	0.41	-
Chew and Lang (1990)	108	4	12	0.41	-
Chew and Morris (1989)	223	5	12	0.60	-
Clancy and Pianta (1993)	193	5	18	0.71	-
Dickinson and Tabors (1991)	62	4	12	0.08	-
Dunn et al. (1995)	46	4	12	0.23	-
Ellwein et al. (1991)	53	4	12	0.44	-
Eno and Woehlke (1995)	161	4	18	0.72	-
Fisher and Fagot (1996)	125	5	24	0.40	-
Funk et al. (1986)	110	4	12	0.64	-
Funk et al. (1986)	118	4	12	0.60	-
Gordon (1988)	109	6	36	0.72	-
Graue and Shepard (1989)	99	5	12	0.24	-
Guerin and Gottfried (1987)	100	5	12	0.37	-
Guerin and Gottfried (1987)	109	5	12	0.16	-
Howes (1990)	80	4	12	0.19	-
Huba et al. (1995)	22	4	24	0.39	-

Table 4 (continued)
Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
La Paro and Pianta (2000) (continued)					
Ironsmith and Poteat (1990)	24	4	24	0.53	-
Jacob et al. (1988)	311	4	12	0.59	-
Jordan (1985)	50	4	18	0.22	-
Jordan (1985)	50	5	18	0.24	-
Kaplan (1993)	50	4	36	0.54	-
Kelly and Peverly (1992)	60	4	18	0.29	-
Kelly and Peverly (1992)	109	5	18	0.49	-
Kontos (1988)	42	4	12	0.51	-
Kontos (1988)	16	5	12	0.38	-
Ladd and Price (1987)	53	4	24	0.25	-
Lowe et al. (1987)	159	5	12	0.51	-
Majsterok and Ellenwood (1995)	79	4	18	0.34	-
Majsterok and Ellenwood (1995)	76	5	12	0.38	-
Mantzicopoulos and Morrison (1994)	232	5	24	0.39	-
Mardell et al. (1990)	42	4	24	0.20	-
Martin et al. (1988)	71	5	12	0.38	-
Martin et al. (1988)	22	4	24	0.23	-
May (1986)	42	5	12	0.75	-
May (1986)	44	5	12	0.61	-
May (1986)	57	5	12	0.78	-
McCormick et al. (1994)	38	5	12	0.41	-
McDevitt et al. (1987)	47	4	18	0.66	-
Pellegrini (1992)	24	5	12	0.35	-
Pettit et al. (1991)	30	4	12	0.22	-
Pianta and Caldwell (1990)	256	5	12	0.49	-
Pianta and Castaldi (1989)	256	5	12	0.24	-
Pianta and Nimetz (1991)	49	5	18	0.31	-
Pianta et al. (1991)	252	5	12	0.30	-
Pilkington et al. (1988)	246	4	24	0.42	-
Porwancher and De Lisi (1993)	119	4	12	0.15	-
Reynolds (1991)	866	5	18	0.47	-
Rose and Wallace (1985)	9	4	24	0.87	-
Rose et al. (1989)	41	4	12	0.69	-
Rose et al. (1991)	59	4	12	0.60	-
Roth et al. (1993)	161	5	24	0.12	-
Schellinger et al. (1992)	26	5	12	0.77	-
Schellinger et al. (1992)	28	4	12	0.54	-
Schmidt and Perino (1985)	378	5	36	0.45	-
Sears and Koegh (1993)	104	5	12	0.57	-
Snider (1997)	50	5	24	0.34	-
Snow et al. (1995)	63	5	12	0.58	-

Table 4 (continued)

Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
La Paro and Pianta (2000) (continued)					
Solan and Mozlin (1986)	40	5	12	0.46	-
Srouge et al. (1990)	164	4	12	0.19	-
Stevenson and Newman (1986)	154	4	36	0.60	-
Stone and Gridley (1991)	176	4	12	0.65	-
Sturner et al. (1996)	343	4	12	0.47	-
Tollefson et al. (1985)	280	5	12	0.60	-
Tsushima and Stoddard (1986)	58	4	24	0.12	-
Uhry (1993)	129	5	12	0.53	-
Vacc et al. (1987)	245	4	24	0.67	-
Wagner et al. (1994)	244	5	18	0.70	-
Walker et al. (1994)	32	4	18	0.41	-
Zucker and Riordan (1990)	75	4	12	0.56	-
Larzelere et al. (2018)					
Barnes et al. (2013)	550	4.1	15	0.48	0.16
Baumrind et al. (2010)	87	4.5	54	0.29	-0.10
Berlin et al. (2009)	2573	2.1	12	0.52	0.04
Coley et al. (2014)	581	3.4	18	0.33	-0.12
Ellison et al. (2011)	456	3	60	0.25	0.09
Fragile Families Study	3575	3	36	0.55	0.08
Gershoff et al. (2012)	11044	6.2	36	0.51	0.08
Gunnoe and Mariner (1977)	1112	7.8	72	0.12	0.13
Lansford et al. (2012)	585	5	12	0.50	0.12
Lazelere et al. (2010)	1464	4.9	12	0.51	0.07
Mendez et al. (2016)	218	2.5	12	0.60	0.03
Mulvaney and Mebert (2007)	979	3	18	0.57	0.09
Straus et al. (1997)	785	7.5	24	0.49	0.07
Pouwles et al. (2016)					
Bagwell and Schmidt (2011)	595	9.2	6	0.44	-
Hodges and Perry (1999)	173	11.3	12	0.84	-
Houbre et al. (2010)	524	9.9	6	0.54	-
Jutengren et al. (2011)	880	13.7	12	0.32	-
Kawabata et al. (2010)	124	9.4	6	0.67	-
Khatri et al. (2000)	471	11.4	12	0.68	-
Malti et al. (2010)	175	6.1	12	0.47	-
Martin et al. (2008)	417	13	12	0.51	-
Monks et al. (2003)	102	5	4	0.78	-
Ostrov (2008)	120	3.7	5	0.42	-
Polasky (2010)	357	9.3	6	0.59	-
Rancourt and Prinstein (2010)	576	12.1	12	0.73	-
Rueger et al. (2011)	694	12.6	6	0.51	-
Salmivalli et al. (1998)	189	14.5	34	0.28	-
Toner and Heaven (2005)	82	12.5	24	0.49	-

Table 4 (continued)
Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Pouwles et al. (2016) (continued)					
Topper er al. (2011)	324	13.9	12	0.21	-
Troop-Gordon and Kopp (2011)	311	10.3	6	0.66	-
Yeung and Leadbeater (2007)	140	9.9	4	0.48	-
Reijntjes et al. (2010)					
Boivin et al. (1995)	641	10.8	12	-	0.09
Bond et al. (2001)	2559	13.5	12	-	0.21
Dhami et al. (2005)	423	6.3	6	-	0.06
Fekkes et al. (2006)	1118	10	7	-	0.26
Goodman et al. (2001)	361	11.5	24	-	0.39
Hanish and Guerra (2000)	1068	7.3	24	-	-0.05
Hanish and Guerra (2002)	1469	7.3	24	-	0.07
Hanish et al. (2004)	126	4.4	6	-	0.10
Hodges and Perry (1999)	173	11.3	12	-	0.23
Hodges et al. (1999)	393	10.7	12	-	0.23
Khatri et al. (2000)	471	11.5	12	-	0.07
Kim et al. (2006)	1666	13.5	10	-	0.05
Kochenderfer and Ladd (1996)	200	5.5	6	-	0.29
Schwartz et al. (1998)	330	9	24	-	0.03
Schwartz et al. (2005)	199	9	12	-	0.41
Snyder et al. (2003)	266	5.5	18	-	0.08
Storch et al. (2005)	144	13.9	12	-	0.18
Sweeting et al. (2006)	2371	11	24	-	0.19
Reijntjes et al. (2011)					
Dhami et al. (2005)	423	6.3	6	-	-0.01
Hanish and Guerra (2000)	1068	7.3	24	-	0.24
Hanish and Guerra (2002)	1469	7.3	24	-	0.21
Hanish et al. (2004)	126	4.4	6	-	0.25
Hodges, Perry (1999)	173	11.3	12	-	0.05
Hodges et al. (1999)	393	10.7	12	-	0.16
Khatri et al. (2000)	471	11.5	12	-	0.10
Kim et al. (2006)	1666	13.5	10	-	0.29
Kochenderfer-Ladd (2003)	379	5.9	12	-	0.12
Ladd and Burgess (2001)	396	5.3	20	-	0.06
Lamarche et al. (2007)	479	6	12	-	0.14
Rusby et al. (2005)	182	10.9	18	-	0.20
Schwartz et al. (1998)	330	9	24	-	0.09
Snyder et al. (2003)	266	5.5	18	-	0.05
Riglin et al. (2014)					
Ansary et al. (2012)	595	11.5	24	-	0.06
Birchwood and Daley (2012)	324	15.5	6	-	-0.26
Capaldi and Stoolmiller (1999)	201	12.5	60	-	0.29

Table 4 (continued)
Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Riglin et al. (2014) (continued)					
Chalita et al. (2012)	237	13.5	12	-	-0.05
Chen et al. (2003)	147	11.5	24	-	-0.08
Fergusseon and Woodward (2002)	964	15.5	60	-	0.19
Flook et al. (2005)	188	10.5	12	-	-0.10
Fredricks and Eccles (2008)	903	12.5	48	-	-0.06
Fredricks and Eccles (2010)	727	13.5	36	-	0.11
Giaconia et al. (2001)	344	18	36	-	-0.19
Gore et al. (2001)	1036	15.5	72	-	0.11
Janosz et al. (2008)	1104	13	12	-	0.03
Kandel and Davies (1986)	924	16	108	-	0.05
Laursen et al. (2002)	181	8	72	-	0.08
Luthar (1995)	138	15	6	-	-0.15
McLeod and Kaiser (2004)	424	15	72	-	-0.08
Miech et al. (1999)	942	15	72	-	-0.32
Morin (2011)	989	12.5	68	-	-0.05
Needham et al. (2004)	10998	14.5	12	-	0.11
Owens et al. (2008)	894	14.5	60	-	-0.16
Schwartz et al. (2005)	199	9	12	-	0.18
Sharma (2005)	700	16	4	-	0.39
Smokowski et al. (2004)	801	16	72	-	-0.05
Suldo et al. (2011)	300	12.5	12	-	-0.08
van Oort et al. (2007)	654	13	120	-	0.19
Vander Stoep et al. (2002)	174	15	30	-	-0.1
Smith et al. (2016)					
Bekes et al. (2015)	47	45.5	50.9	0.38	0.17
Dunkley et al. (2006)	96	34.3	158.6	0.20	0.27
Dunkley et al. (2009)	107	34.4	192	0.21	0.28
Enns et al. (2001)	96	25.1	24	0.22	-0.07
Enns et al. (2005)	206	24	20	0.44	0.15
Graham et al. (2010)	240	20	3	0.60	0.08
Mackinnon and Sherry (2012)	127	18.3	19	0.47	0.21
Mackinnon et al. (2012)	226	22.4	4	0.57	0.04
Mackinnon et al. (2012)	226	21.5	4	0.60	0.08
Sherry et al. (2013)	155	20.7	4.3	0.40	0.13
Sherry et al. (2014)	232	50.1	3.3	0.52	0.12
Smith et al. (2018)					
Akram et al. (2015)	76	25.3	24	0.81	0.07
Damian et al. (2017)	489	15	9	0.59	0.11
Einstein et al. (2000)	508	17.6	4	0.55	0.06
Flaxman et al. (2012)	77	46	4	0.43	0.25
Herman et al. (2013)	547	6.22	144	0.18	0.01

Table 4 (continued)

Included Studies

Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Smith et al. (2018) (continued)					
Joiner and Schmidt (1995)	174	19.8	1.5	0.49	-0.12
Mandel et al. (2015)	150	41.02	96	0.49	0.13
O'Connor et al. (2010)	515	15.2	12	0.61	0.07
Oddo-Sommerfeld et al. (2016)	266	32.35	9	0.38	0.13
Sherry et al. (2014)	302	20.84	6	0.67	0.02
Valentine et al. (2004)					
Anderman (1999)	312	12.5	12	-	0.13
Bachman (1986)	1487	15.5	8	-	0.10
Boehm-Morelli (1999)	106	8.5	2	-	0.01
Bradley (2000)	503	-	5	-	0.10
Brudos (1995)	206	9.5	36	-	0.18
Chan (In Press)	25	0.93	9	-	0.08
Chapman (1981)	166	11	5	-	0.13
Chapman (1981)	208	9	5	-	-0.01
Chapman (1988)	77	11.3	9	-	0.29
Chapman (1988)	70	11.3	9	-	0.34
Chapman (1997)	117	5.1	12	-	0.09
Chemers (2000)	256	19	5	-	0.36
Coon-Carty (1998)	73	9.5	7	-	0.23
Cross (2001)	123	18.5	20	-	0.11
DuBois (1999)	332	11.5	24	-	0.08
DuBois (1999)	144	13.4	12	-	0.05
Entwisle (1987)	155	6.5	18	-	0.11
Entwisle (1987)	130	6.5	18	-	0.08
Entwisle (1987)	162	6.5	18	-	0.04
Entwisle (1987)	129	6.5	18	-	0.04
Geroski (1996)	141	10.5	2	-	0.19
Guay (1999)	396	9	12	-	0.19
Helmke (1995)	696	7.5	12	-	0.13
Hemsley (1991)	217	13.5	24	-	0.14
Hemsley (1991)	98	13.5	24	-	0.01
Hemsley (1991)	69	13.5	24	-	0.05
Kong (2000)	5985	13.5	12	-	0.05
Kurtz-Costes (1994)	45	8.5	24	-	0.13
Marsh (1988)	107	7.4	6	-	0.10
Marsh (1991)	6777	15.8	24	-	0.10
Marsh (1997)	402	12.7	2	-	0.18
Marsh (1998)	6002	13.5	4	-	0.01
Marsh (1999)	927	15.9	2	-	0.01
Marsh (2000)	7990	13.5	12	-	0.04
Maruyama (1981)	145	12	6	-	0.02
Maruyama (1981)	159	12	48	-	0.06

Table 4 (continued)

Included Studies

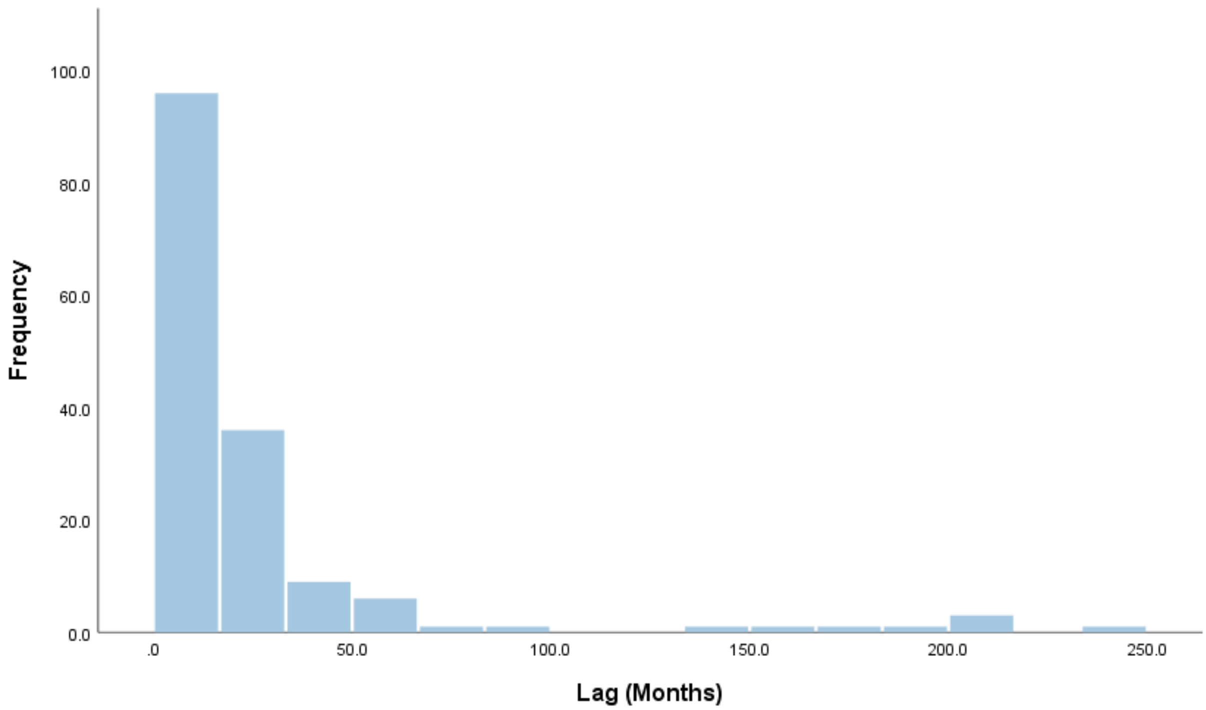
Author (Year)	N	M_{Age} (Time 1)	Lag (Months)	Stability	Cross-lag
Valentine et al. (2004) (continued)					
Mijus (1997)	889	9.5	12	-	0.11
Mone (1995)	214	-	2	-	0.02
Mundy (2000)	37	-	12	-	0.10
Sharrow (1993)	59	13.5	12	-	0.02
Shavelson (1982)	99	14	4	-	0.11
Shoemaker (1980)	244	10.5	36	-	-0.10
Simmons (1987)	276	11.5	17	-	0.16
Skaalvik (1990)	363	11.5	12	-	0.00
Skaalvik (1990)	363	8.5	12	-	0.04
Skaalvik (1990)	493	11.5	12	-	0.07
Skaalvik (1999)	493	8.5	12	-	0.04
Skaalvik (1999)	225	13.5	12	-	0.11
Thordardottir (2000)	106	9.5	7	-	0.17
Thordardottir (2000)	107	12.5	7	-	0.15
Thordardottir (2000)	121	15.5	7	-	0.03
Van Damme (2000)	6410	6	12	-	0.10
Widlak (1983)	83	7.5	6	-	0.07
Williams (1998)	141	6.5	24	-	0.17
Yeung (1999)	485	13.5	3	-	0.00
Yin (1999)	542	20	6	-	0.01
Yoon (1996)	462	11.5	36	-	-0.12
Yoon (1996)	362	11.5	36	-	-0.01
Zimmerman (1992)	101	15	6	-	0.08
Zimmerman (1997)	1057	11.5	12	-	0.08
Wardle et al. (2011)					
Brunner (2007)	2594	-	228	-	0.04
Brunner (2007)	1060	-	228	-	0.01
Fowler-Brown (2009)	272	-	156	-	0.00
Fowler-Brown (2009)	474	-	156	-	0.08
Gerace (1996)	438	-	84	-	0.10
Helminen (1999)	133	-	12	-	0.00
Ishizaki (2008)	2200	-	72	-	0.03
Ishizaki (2008)	1371	-	72	-	0.02
Lallukka (2008)	228	-	336	-	0.00
Lallukka (2008)	155	-	336	-	0.00
Lloyd (1996)	301	-	24	-	0.08
Lloyd (1996)	291	-	24	-	0.10
Rookus (1988)	193	-	24	-	0.01
Rookus (1988)	250	-	24	-	0.00
Thomas (2008)	9219	7	456	-	0.00
Twisk (1999)	166	27	24	-	-0.05
Vitaliano (1996)	54	-	16.5	-	0.00

Note. The above table shows non-transformed stability and cross-lagged coefficients. M_{Age} refers to participant mean age at Time 1.

Appendix C

Figure 6

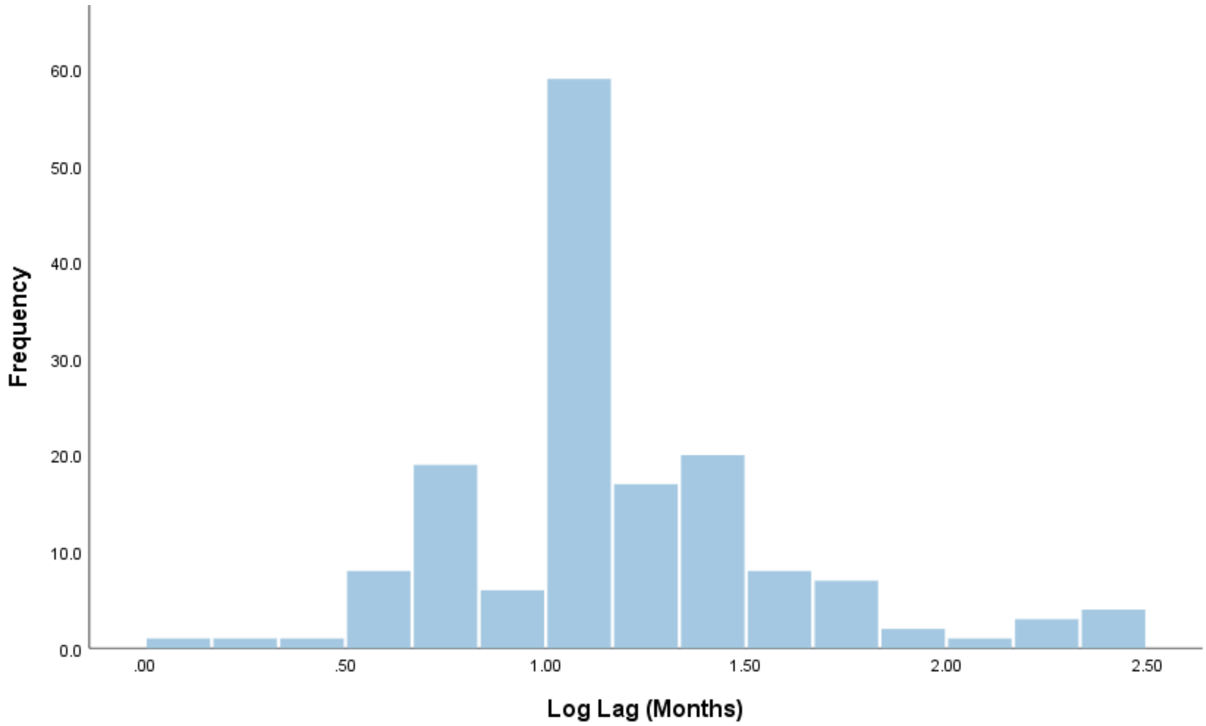
Bar Graph of Untransformed Lag Lengths: Stability



Note. Lag lengths not transformed.

Figure 7

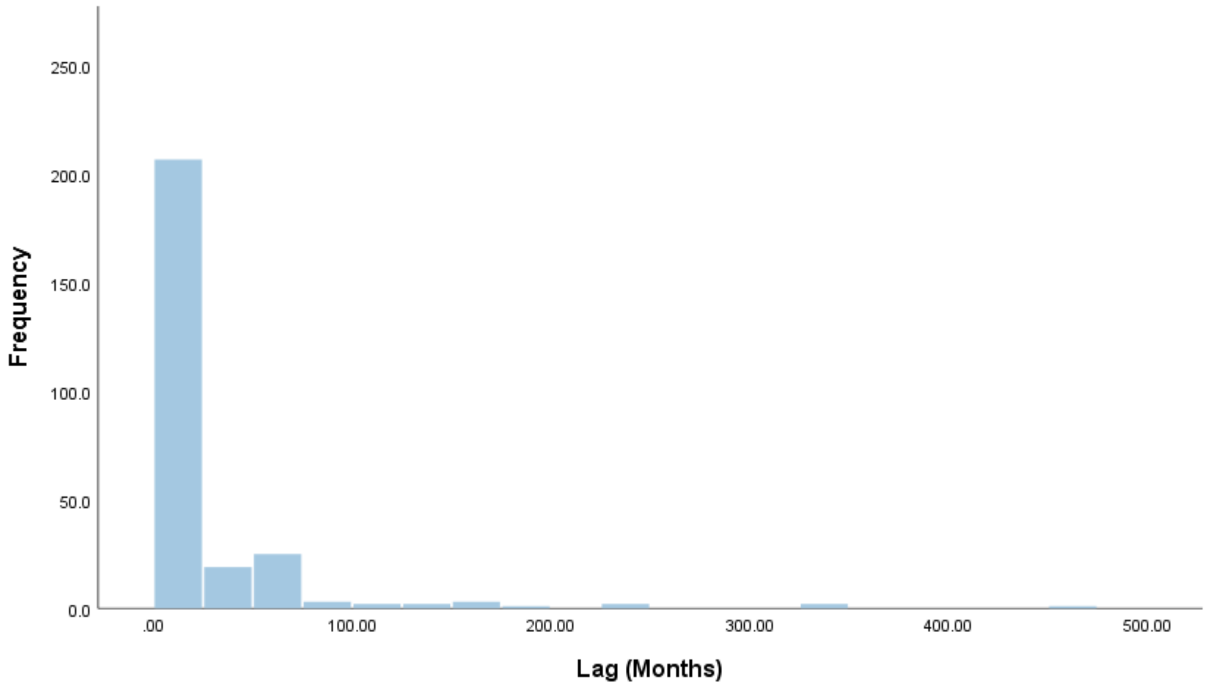
Bar Graph of Transformed Lag Lengths: Stability



Note. Vertical axis represents how frequently a lag length was used, log transformed

Figure 8

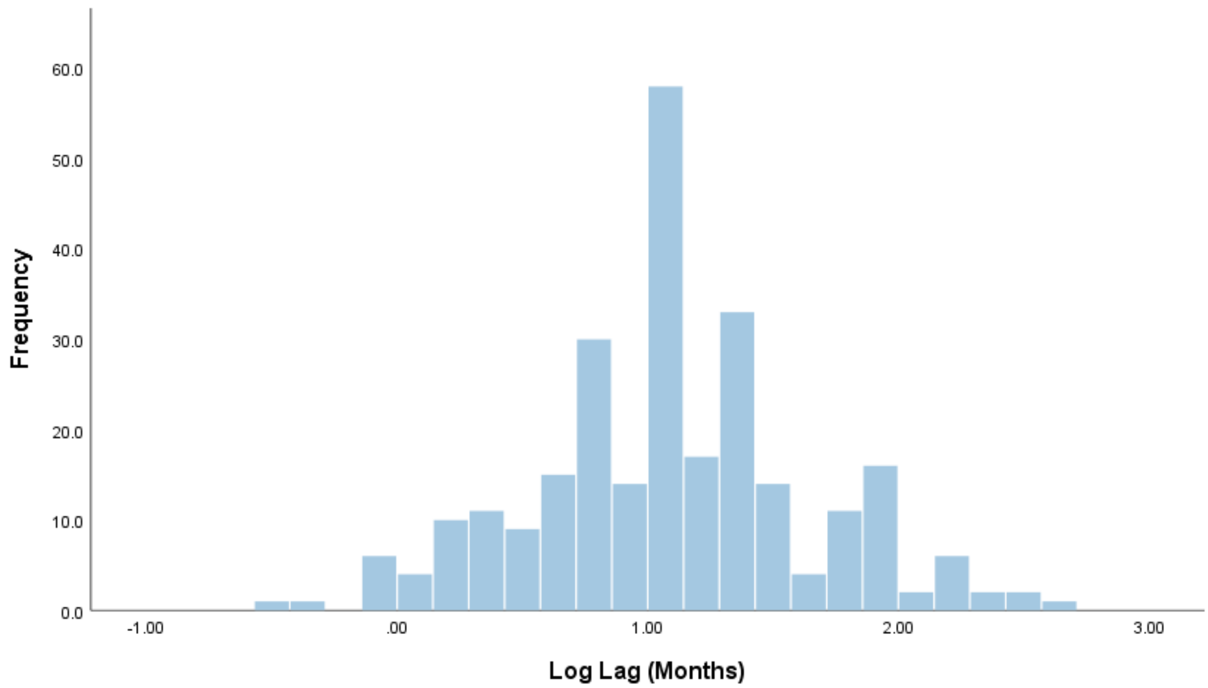
Bar Graph of Untransformed Lag Lengths: Cross-lag



Note. Lag lengths not transformed.

Figure 9

Bar Graph of Transformed Lag Lengths: Cross-lag

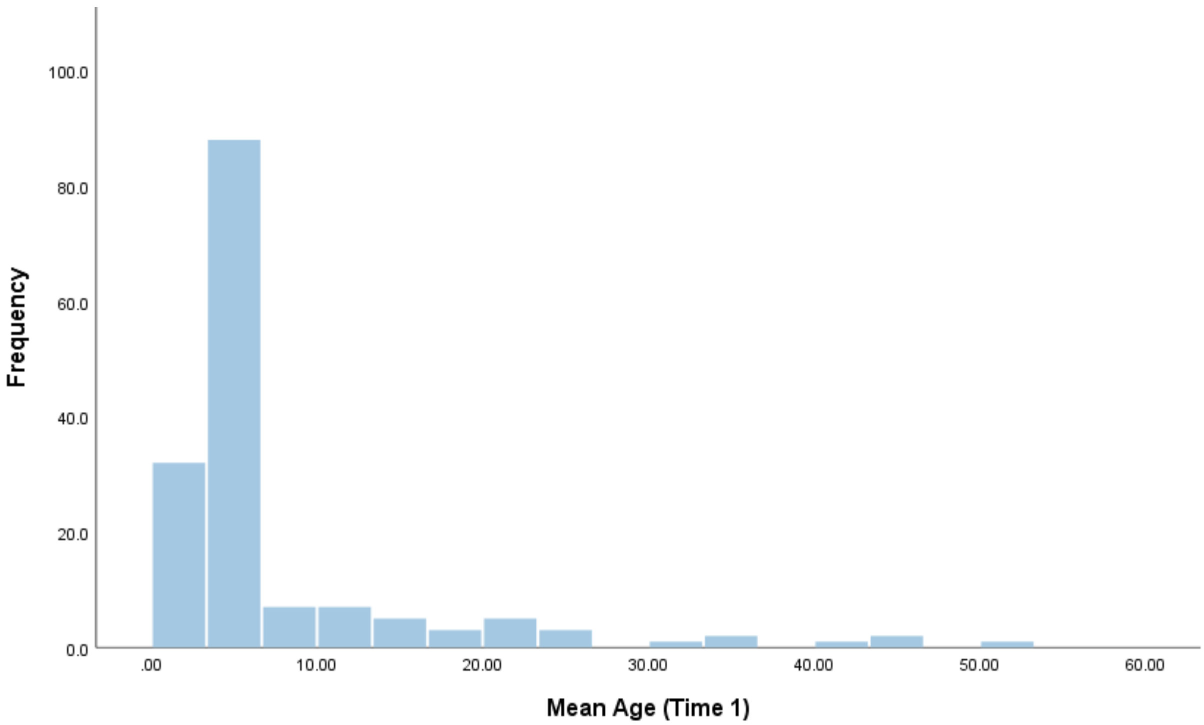


Note. Vertical axis represents how frequently a lag length was used, log transformed.

Appendix D

Figure 10

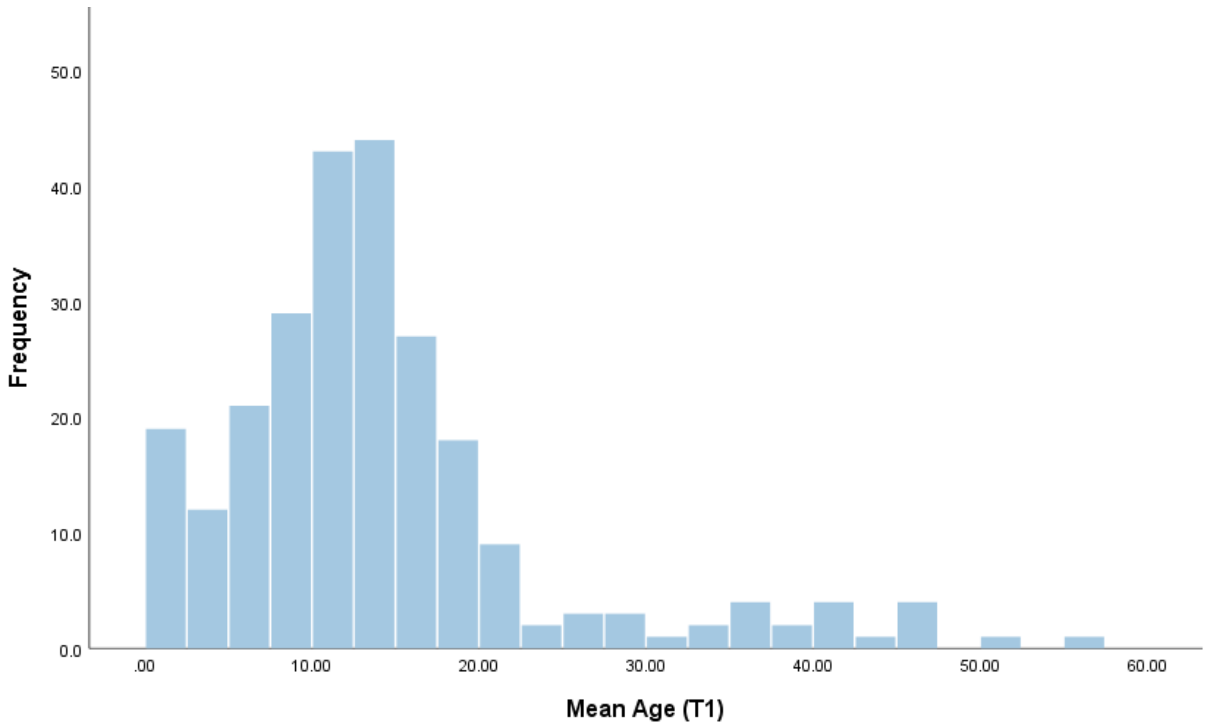
Bar Graph of Age for Stability



Note. The distribution of age is positively skewed, indicating that many more participants were from younger age groups than older ones.

Figure 11

Bar Graph of Age for Cross-lag

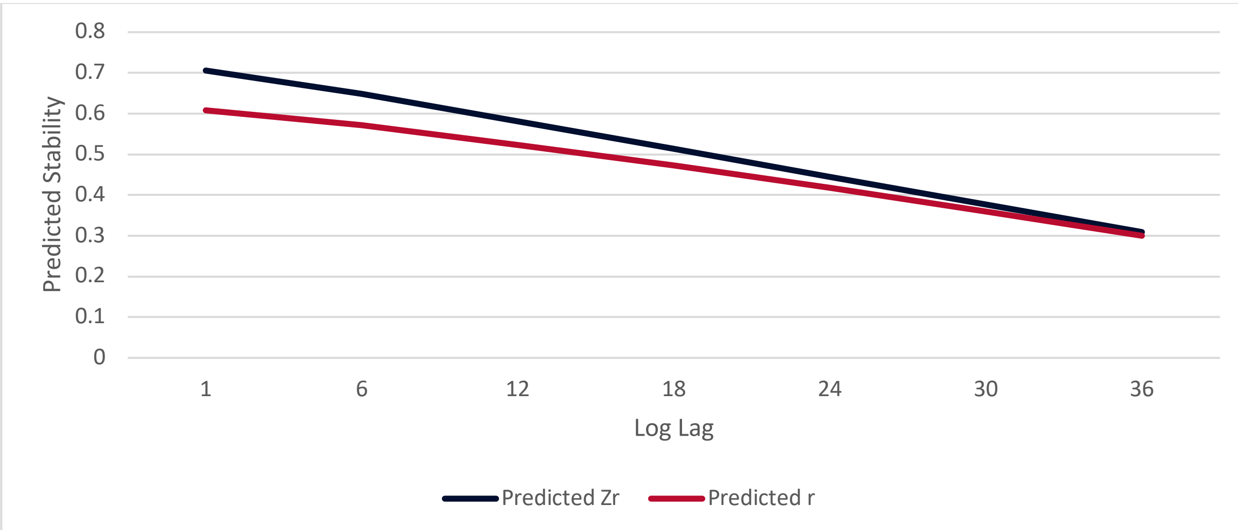


Note. The distribution of age is positively skewed, indicating that many more participants were from younger age groups than older ones.

Appendix E

Figure 11

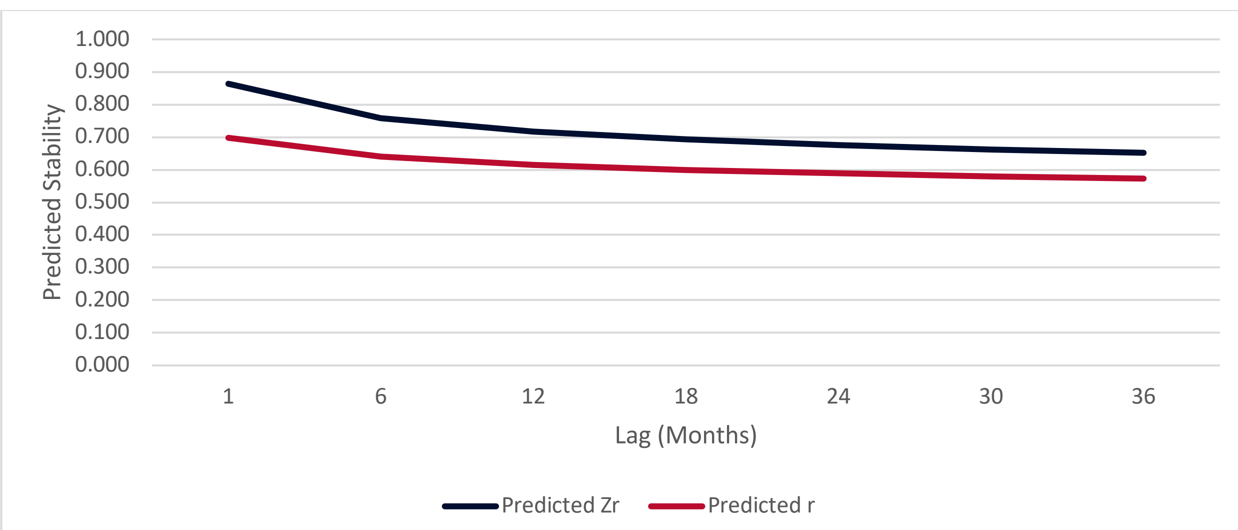
Predicted Linear Decay in Stability Using Log-transformed Lag



Note. The blue line represents predicted stability (*Zr*) and the red represents back-transformed Pearson's *r*.

Figure 12

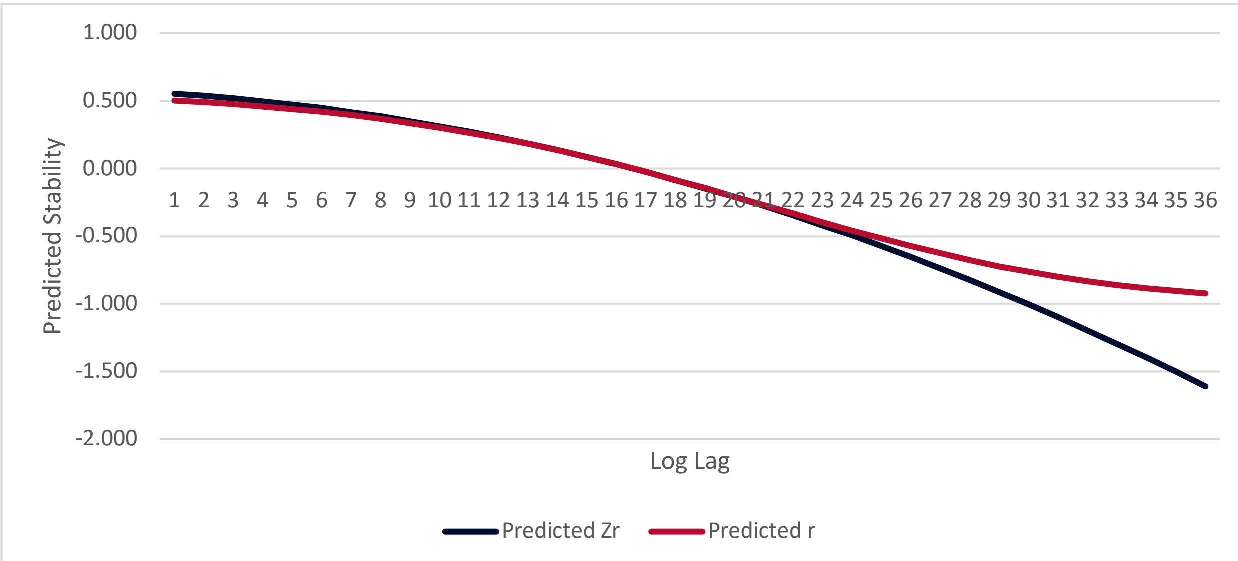
Predicted Linear Decay in Stability Using Back-transformed Lag



Note. The blue line represents predicted stability (*Zr*) and the red represents back-transformed Pearson's *r.*

Figure 13

Predicted Quadratic Decay in Stability Using Log-transformed Lag

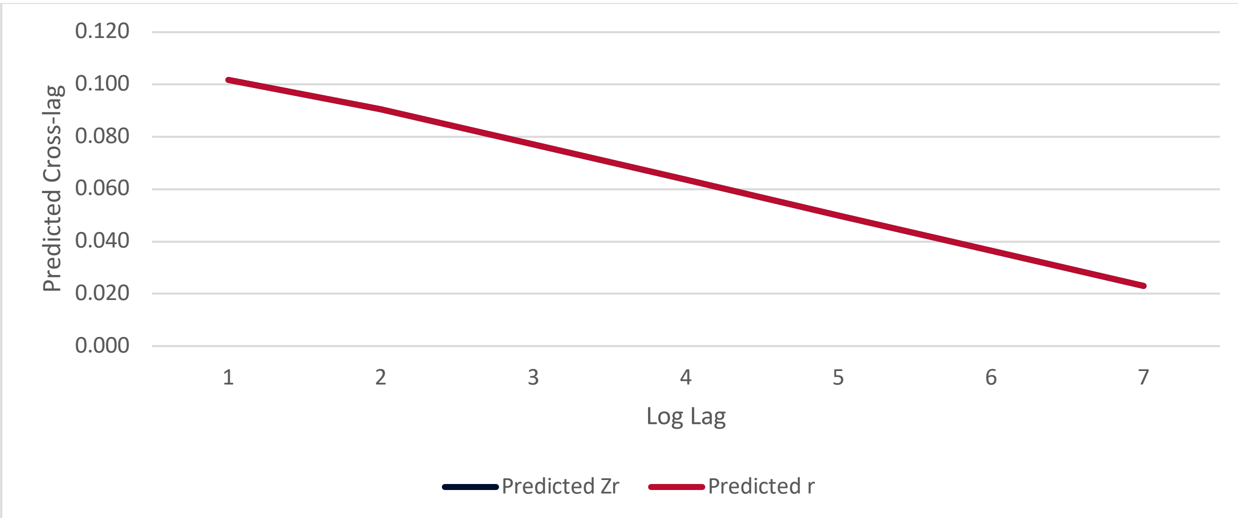


Note. The blue line represents predicted stability (*Zr*) and the red represents back-transformed Pearson's *r*.

Appendix F

Figure 14

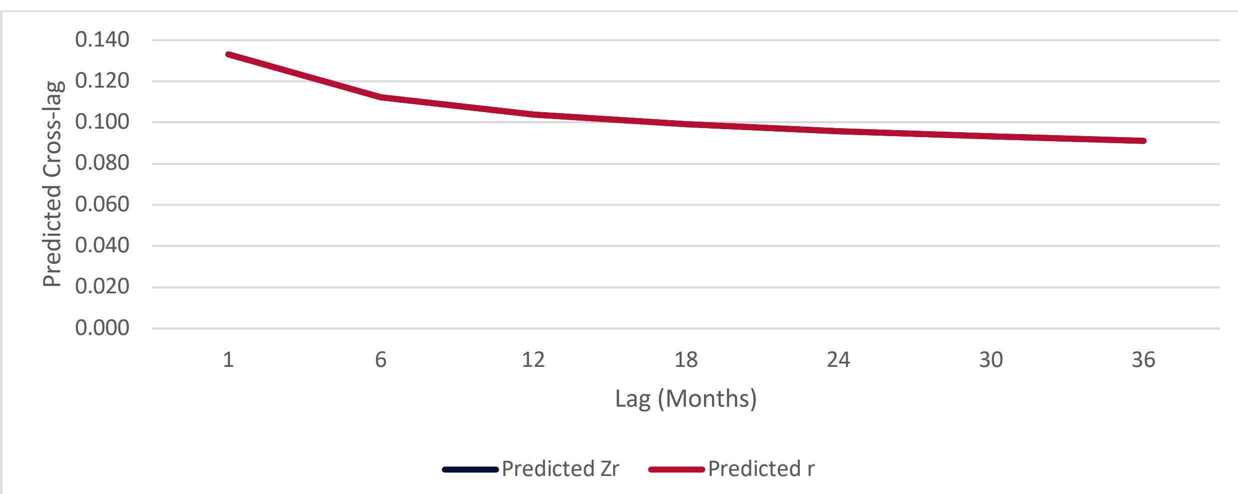
Predicted Linear Decay in Cross-lag Using Log-transformed Lag



Note. The blue line represents predicted stability (*Zr*) and the red represents back-transformed Pearson's *r*. In the above graph, Predicted *Zr* and *r* overlap. The difference in conversion exists at higher and lower values those shown within the range predicted above.

Figure 15

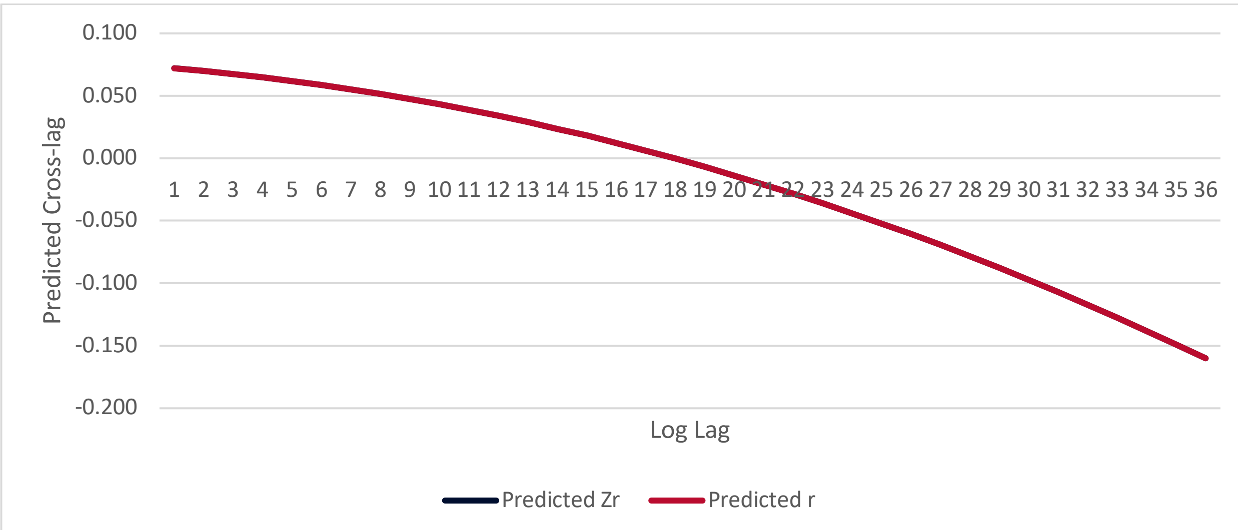
Predicted Linear Decay in Cross-lag Using Back-transformed Lag



Note. The blue line represents predicted stability (*Zr*) and the red represents back-transformed Pearson's *r.* In the above graph, Predicted *Zr* and *r* overlap. The difference in conversion exists at higher and lower values those shown within the range predicted above.

Figure 16

Predicted Quadratic Decay in Cross-lag Using Log-transformed Lag



Note. The blue line represents predicted stability (*Zr*) and the red represents back-transformed Pearson's *r.* In the above graph, Predicted *Zr* and *r* overlap. The difference in conversion exists at higher and lower values those shown within the range predicted above.