Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Emily Hannan Beattie April 8, 2020

Understanding Variation in Self-Derivation Performance: Individual and Study-Level Meta-Analyses of Bauer Memory Lab Self-Derivation Studies

by

Emily H. Beattie

Dr. Patricia Bauer

Adviser

Psychology

Dr. Patricia Bauer

Adviser

Dr. Yuk Fai Cheong

Committee Member

Dr. Jin Kim

Committee Member

Dr. Jessica Dugan

Committee Member

2021

Understanding Variation in Self-Derivation Performance: Individual and Study-Level Meta-Analyses of Bauer Memory Lab Self-Derivation Studies

By

Emily H. Beattie

Dr. Patricia Bauer

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Psychology

2021

Abstract

Understanding Variation in Self-Derivation Performance: Individual and Study-Level Meta-Analyses of Bauer Memory Lab Self-Derivation Studies
By Emily H. Beattie

The current research was an examination of individual and study-level factors that explain the variation in performance on derivation of new semantic knowledge through integration of information. To examine these factors across all studies that examine self-derivation in the Bauer Memory Lab at Emory University, I used an individual participant data (IPD) meta-analysis and a traditional study-level meta-analysis. Using the IPD meta-analysis, I used six studies to examine the associations of verbal comprehension, semester (fall or spring) and the half of the semester, age, race, and gender with self-derivation, while controlling for the specific study. In the study-level meta-analysis, I used 11 studies to examine the associations of the presentation of facts modality, number of facts presented to learn, and assessment method with self-derivation. In the IPD meta-analysis, I found that age and verbal comprehension were significantly associated with self-derivation performance across studies and explained a significant portion of the variability in performance. Additionally, in the study-level meta-analysis, I found that none of the predictors explained any of the variability across studies. The present research strengthens previous findings that verbal comprehension is implicated with self-derivation performance by showing that this trend is consistent across multiple studies. In addition, this study provides greater insight into why some people perform vastly better at self-derivation tasks than others. Ultimately, these findings provide important general findings regarding factors implicated in the variation of self-derivation performance and establish future research goals to enhance understanding of semantic learning.

Keywords: self-derivation, variability, verbal comprehension, meta-analysis, Bauer Memory.

Understanding Variation in Self-Derivation Performance: Individual and Study-Level Meta-Analyses of Bauer Memory Lab Self-Derivation Studies

By

Emily H. Beattie

Dr. Patricia Bauer

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Psychology

2021

Table of Contents

Understanding Variation in Self-Derivation Performance: Individual and Study-Level Meta-Analyses of Bauer Memory Lab Self-Derivation Studies

Every day, humans are bombarded with vast amounts of unique information they will either keep in mind or will forget. Whether through formal learning, such as in a classroom, or informal learning, such as reading a newspaper, humans continuously grow their semantic (factual) knowledge base. Humans combine newly learned information with pre-existing knowledge, thereby resulting in an extensive knowledge base. They also go beyond what they have learned directly, through productive processes such as inference. One of the cognitive productive processes that allows for the formation of new semantic knowledge is self-derivation. In the process of self-derivation, separate but related facts are combined to form new information (Bauer & Jackson, 2015; Varga & Bauer, 2017; Varga, Esposito & Bauer, 2019). Without such a process, humans would be limited to only learning information that is presented to them. Yet self-derivation is marked by individual differences, such that some people are more successful at deriving novel facts than others. The purpose of the present research was to utilize past research in an individual participant data meta-analysis and a study-level meta-analysis in order to determine which individual measures and study design factors are associated with performance on self-derivation tasks.

This study utilizes the work of the Bauer Memory Lab at Emory University. This lab studies the process of self-derivation and various factors that are associated with successful derivation. In the standard self-derivation paradigm, a participant is presented with two separate, but related facts (known as stem facts; e.g., “Kierkegaard was the founder of Existentialism” and “The self is the main focus of Existentialism”). By integrating these two facts based on their shared information (e.g., Existentialism), a novel fact, “Kierkegaard developed a philosophy that

emphasized the self" can be formed. (Dugan & Bauer, 2020). When adults are presented with only one of these stem facts, they do not successfully self-derive new information (Varga & Bauer, 2017).

When studying self-derivation, it is critical to not only study the process, but the variation in self-derivation of novel facts. For example, in studies examining self-derivation, there is consistently a wide range in performance. Specifically, correct scores have varied from 3% to 93% (Varga & Bauer, 2017b; Varga, Esposito & Bauer, 2019) and 7% to 100% (Varga & Bauer, 2017a) in adults. Additionally, these wide ranges of approximately 90% are apparent in studies yet to be published by the Bauer Memory Lab. This consistent, broad range in scores raises the question of which factors that are influential in making some people better at self-deriving. Although some studies have examined the associations between self-derivation and individual difference measures of working memory and verbal comprehension (Varga, Esposito & Bauer, 2019), there has not yet been a systematic analysis of these correlates, nor has there been an analysis of study-level factors, such as how facts are presented. It is important to understand what is associated with this variability to find ways to possibly optimize self-derivation, and thus, render them better learners. To fill this gap, in the present study I examined which individual and study-level factors are associated with the variability in self-derivation scores across several studies through a meta-analysis. This method allows us to see whether factors that are associated with success self-derivation are consistent across studies.

A review of the literature suggested several individual and task factors that may contribute to variability in self-derivation performance. The ones of focus in this study will be further discussed below. However, outside of the domain of self-derivation, a vast amount of psychological research has examined variability in performance to explain why some people

perform better than others at certain tasks. Patterns in variability are present across various studies of cognition. For example, at the level of the task, greater performance on the first trial of a screen-searching task is associated with a higher level of peak performance on the same task (Ericson, Kravitz & Mitroff, 2017). At the level of the individual, neuronal research has also been conducted to assess performance. Specifically, hippocampal and midbrain connectivity has been implicated in ability to generalize new information, a process resembling self-derivation (Shohamy & Wagner, 2008). These findings imply that performance is impacted both at the superficial task level and at the individual level. Since variability in performance is of interest in most psychological domains, it is critical to understand what influences this variability to gain greater insight into human performance.

Individual-level measurements

In order to assess individual participant score variation, I examine measurements that exist at the individual level. This query is driven by the findings of Varga, Esposito and Bauer (2019), who studied how various cognitive measures are associated with the variation in self-derivation performance. In this study, verbal comprehension was significantly positively associated with self-derivation performance, $r = 0.53$, and explained unique variance in scores, $R^2 = 35\%$ (Varga, Esposito & Bauer, 2019). Verbal comprehension was assessed using the Woodcock-Johnson III Test 1. This test is meant to measure one's verbal, semantic knowledge (Woodcock, McGrew & Mather, 2001; Varga, Esposito & Bauer, 2019). The test consists of four subtypes: picture vocabulary (identify the name of an object in a picture), synonyms, antonyms, and analogies (e.g., boat is to sea as plane is to ___). The synonym tasks require participants to say words similar in meaning to a word presented, and antonym tasks require

participants to say a word opposite in meaning to what is presented. Upon completing these tasks, a sum is computed for each subsection, as well as an overall score.

Despite the previous findings of a relations between verbal comprehension scores and self-derivation, it has only been examined in one study to date. As a result, it is currently unknown whether the positive association between verbal comprehension and self-derivation is consistent across studies. In many of the studies conducted in the Bauer Memory Lab, the Woodcock-Johnson III Test 1 for verbal comprehension was administered to nearly every participant as a buffer task between encoding (being shown the stem facts) and assessment. As such, it is possible to conduct a meta-analysis using verbal comprehension as a predictor to see whether it remains a significant positive predictor in self-derivation performance.

In addition to verbal comprehension, the half of a semester and the semester (fall or spring) that a student participates in research could be associated with variability in self-derivation performance. Semester has potential relevance in this study because each of the Bauer Memory Lab projects involved participants from a private university. As a result, it is important to consider the part of the semester a trial is conducted. In fact, the end-of-the-term participants have been found to perform worse in research studies than participants from the beginning of the term (Porfido et al., 2020). Students oftentimes are required to enroll in studies to fulfill a requirement for the introductory psychology course. While this trend found by Porfido and colleagues (2019) may reflect procrastination in completing study credit, participants may also be experiencing burnout over the course of the semester. As a result of the multitude of demands on college students during the academic year, many college students experience some form of stress. When students do not engage in stress-reducing activities, perceived stressed, test anxiety, and personal burnout increase over the course of the semester (Baghurst & Kelley, 2014).

Additionally, first semester tends to be associated with greater stress than second semester of an academic year (Maricutoiu & Sulea, 2019). Acknowledging these findings, an increase in student stress may be negatively related with performance in a laboratory setting. If personal burnout is indeed increasing over the course of the semester, then it would be expected that the latter half of the semester would be associated with lower scores of self-derivation performance and would thus be a significant in the explanation of unique variance. Regardless of the reason for poorer performance, there is an ample amount of research that indicates that the end of semester may introduce variation in academic study performance.

Study-level measurements

Although some factors can be examined at the individual level, some measurements and methods must be studied at the study level. Even though each study used the same task, each had a unique goal to examining self-derivation. As a result, many of the studies varied in how they administered tasks to participants. With this idea in mind, study-level factors consist of variables that are held constant within each study. One of these factors is the number of facts participants were given to encode. Some studies used as few as 10 facts total while other used as many as 120 total. As the number of facts presented to participants increases, it can impose a mental burden, especially since all facts are meant to be novel to participants. If any of the facts are not successfully encoded, then the novel stem fact that results of integration cannot be formed (Bauer & Varga, 2017). Perhaps, in studies where participants were given more facts to encode, there is a lower level of self-derivation performance. I theorize that this phenomenon may be the case as a result of attentional overload. Attentional overload is when there are excessive demands on attention, resulting in a depletion in available attentional resources (APA Dictionary, 2020). In many cases, when attentional resources are depleted, performance on tasks declines

(Matthews, Sparkes & Bygrave, 1996). If participants in a study with many facts (e.g., 120) were distracted by the large quantity or were mentally overloaded, perhaps self-derivation performance is lower.

Modality, referring to how facts are presented, is an additional study-level characteristic that may be implicated in the variability in self-derivation performance. In some studies, some facts were presented in a narrative form, consisting of a short paragraph on a topic. In other studies, facts were presented using a power point in sentence format. For this study design, slides were presented to the participants long enough for them to read it before the next fact appeared. How facts were presented may result in participants processing them differently. As such, if one study utilized one modality that is more facilitative to learning, then that study may have a greater effect size in self-derivation performance than a study using a less effective modality. Research outside of the Bauer Memory Lab has examined how words and images result in various levels of successful learning, particularly when examining episodic memory. For example, people have been found to have a greater ability to remember pictures than words (Grady et al., 1998). This trend has also been studied at the neuronal level, where researchers found that there was greater activity in the medial temporal cortex during the encoding of pictures (Grady et al., 1998). The medial temporal cortex is a portion of the brain that is implicated in episodic memory, and greater activity in this area signals that new memories are being stored. If this research holds true for semantic memory, perhaps modalities that include pictures rather than just sentences or narrative would be associated with greater effect sizes in self-derivation performance.

Finally, the assessment methods of successful self-derivation may be implicated in performance. Generally, participants were assessed on their self-derivation performance using

open-ended and forced-choice questions. Open-ended involved participants writing the linking answer to a prompt (e.g., Kierkegaard developed a philosophy that emphasized _____), whereas forced-choice involved participants selecting an answer in a multiple-choice format. Open-ended assessment tends to be more challenging for participants, since they must successfully derive the fact independently without looking at possible options (Esposito & Bauer, 2018). Additionally, open-ended prompts are generally the primary testing form used because forced-choice questions provide answer options that may trigger the remembrance of facts that were presented. If forced-choice questions were given instead of open-ended questions, then general performance of all participants would likely improve since participants do not have to retrieve the related information independently (Glass & Sinha, 2013; Esposito & Bauer, 2018). In the current standard assessment method for the Bauer Memory Lab, the open-ended integration questions are provided first, followed by the forced-choice integration questions. This method was developed to ensure that participants first are assessed in their ability of deriving the novel fact independently, then it is determined if they could derive with options. As a result of this method, the first form of assessment offered to participants in a study is used to calculate the self-derivation score. This meta-analysis includes assessment method in order to see if performance on forced-choice integration is better than performance on open-ended integration across all studies.

The current research

Though research has been conducted on the relations between cognitive measures and self-derivation (Varga, Esposito & Bauer, 2019), no research has been conducted examining whether the patterns are consistent across studies. To best assess variation in self-derivation, I conducted two meta-analyses: an Individual Participant Data (IPD) meta-analysis and a

traditional study level meta-analysis. An IPD meta-analysis allowed me to examine the factors in relation to each individual participant across the included studies. A traditional meta-analysis then allowed me to examine study-level factors that vary across study, but not participant.

I chose to use a meta-analysis to address the question presented in this study because it is a quantitative approach. A meta-analysis is a quantitative form of study design that is used to assess multiple studies to make conclusions regarding a topic (Tavakol, 2019). Often, meta-analyses are used in order to increase statistical power and to answer questions that are not directly asked in the individual studies (Hoffman, 2015). Since this study was conducted by the Bauer Memory Lab using their existing data, it was possible to extract effect sizes and to analyze each individual participant included. Another option was to conduct a literature review, which involves discussing an assortment of studies in conjunction with one another to synthesize new information (*Writing a Literature Review*, 2020). A meta-analysis builds on a literature review because the synthesis is supported by quantitative and statistical measures. As a result, I found a meta-analysis to be the best analytic strategy for this empirical study. By using the meta-analysis approach, I will be able to gain more insight into which individual and study factors are associated with the consistent range of scores across studies.

The IPD meta-analysis (Study 1) examined verbal comprehension, semester half and season, in addition to age, gender, and race as controlling factors, in relation to self-derivation. Based on prior research, I predict that verbal comprehension and semester half will be associated with self-derivation, with higher verbal comprehension scores and first semester half being associated with greater self-derivation performance. The traditional meta-analysis (Study 2) examined the association between number of facts, assessment method, and modality with self-derivation scores. Based on the previous findings, I predict that these variables will be associated

with self-derivation and explain a portion of the variability in self-derivation across studies. These findings could provide more insight on what makes some people efficient integrators and if there are some modifications that can be made in learning environments to facilitate one's ability to learn. Additionally, the findings can provide more insight into the variation of self-derivation scores, specifically, why the variation is consistent across each of the studies. In total, I examined 11 studies, including 543 participants. The work is expected to contribute to the literature by providing greater insight into which factors optimize integration and self-derivation performance. With this information, we can further explore ways to make learning more effective.

Study 1

Methods

Search Strategy and Study Selection

The search for potential studies was conducted by accessing the server of the Bauer Memory Lab at Emory University in Atlanta, Georgia. This server houses the entirety of research conducted by the lab with individual participant data, protocol, and analyses. Each study on the server was examined to see if it adhered to the following criteria:

(1) The study examined the standard self-derivation paradigm, meaning participants were provided with stem fact pairs and then assessed on their ability to form a novel fact,

(2) The study was conducted at the Bauer Memory Lab at Emory University. This criterion was included since, to date, all studies of self-derivation have been conducted at the Bauer Memory Lab.

(3) The study was conducted on an adult population (university level or higher).

(4) Demographic information, such as gender, age, and race, was collected from participants.

(5) Participants were administered the Woodcock Johnson III - Test 1 of Verbal Comprehension

Using the above criteria, six studies were selected for inclusion in this study out of the 110 available studies. In total, one study is published (200), one is under review (206), one has been accepted for publication (248), and three are as yet unpublished (189, 231, 253). Each study is denoted by its numeric value on the server. Additionally, prior to examining any descriptive statistics, I preregistered the methods of this study and my hypotheses with As.Predicted.org.

Participants

Upon compiling the data from each study, 340 participants were included (M = 19.55 years, SD = 1.12 years, 222 females) from the six studies. After examining the demographics reported for each study, the compiled list of participants from each study consisted of 0.3% American Indian or Alaskan Native, 26.2% Asian, 12.1% Black or African American, 55.8% White or Caucasian, 4.7% Multi-racial, and 0.8% did not report. Each of the projects included in this study had received IRB approval when they were conducted. Participants earned course credit for their participation.

Study Design and Outcomes

Data files, master participant lists, and protocol were extracted from each study and organized for analysis. During this compiling process, if a certain measure used in one study was not utilized in another study, then those participants were given NA values for that measure. The following variables were extracted for analysis:

Study Number. The study the participants engaged in was noted, allowing participants to be clustered by study (see statistical analysis section for further explanation).

Spring or Fall Semester. The academic semester in which the participant completed the protocol was denoted as either being Fall or Spring. If a study was conducted between August and December, it was assigned the value of "fall." If a study was conducted between January and May, it was assigned the value of "spring." In a study, four participants were conducted in the summer months. As such, they were assigned the value of "summer."

First or Second Half Semester. The date of participation was denoted as being either during the first half of the academic semester or second half of the academic semester. Based on the academic calendar of the institution, the average semester was found to be approximately 150 days. As a result, October 24 was used as the half-way point during the fall semester and March 14 was used as the half-way point during the spring semester. This variable is meant to assess whether students perform better with the self-derivation task at the beginning or the end of the semester.

Age. This variable was defined as the age of the participant at the time the study was conducted, in years; months.

Gender. This variable was defined as the sex of the participant, coded as either "female" or "male."

Race. Participant race was coded as "American Indian or Alaskan Native," "Native Hawaiian or Pacific Islander," "Black or African American," "Asian," "White or Caucasian," "Multi-Race," or "did not report."

Verbal Comprehension. Verbal comprehension was assessed using the Woodcock-Johnson III - Test 1. As discussed in the introduction, this test measures verbal, semantic knowledge.

Proportion of Self-Derivation Correct. The proportion of successful self-derivation is the outcome of interest in this study (e.g., answering "the self" to the prompt "Kierkegaard developed a philosophy that emphasized ___?___" is a successful self-derivation). This value represents the proportion of self-derivation assessment questions that the participant answered correctly. In most cases, this value was calculated from the open-ended integration questions. However, Study 181 only used forced-choice integration questions. As a result, the value for these participants were calculated using the forced-choice integration responses.

Statistical Analysis

A meta-analysis of individual participant data (IPD) was conducted to assess the impact of the variables above on self-derivation. IPD meta-analysis allows for the incorporation of unpublished data and for the examination of self-derivation at the individual level and the study level (Riley et al., 2020).

To examine the relation between the selected potentially explanatory variables and successful self-derivation, a multiple regression model was created. A fixed effects model was used to examine whether a significant relation existed between each of the predictors and self-derivation. I included participant age, gender, and race in the multiple regression model to control for possible effects of these demographic variables on predictions of self-derivation. Additionally, the study number was included in the analysis to control for the effects of different studies. This form of clustering is necessary because analyzing the data as if it is from a single study could result in significant differences not being detected (Riley et al., 2020). Following

identifying significant predictors, a regression was conducted with these significant predictors to examine the coefficient of determination (R^2). This value indicates what percentage of the variability in self-derivation can be explained by the predictor variables.

Results

For this study, I aimed to answer which individual-level variables predict self-derivation performance. In order to address this question, I conducted an individual participant data (IPD) meta-analysis using a multiple regression linear model. The individual-level variables include verbal comprehension scores on the Woodcock Johnson III (Woodcock, McGrew & Mather, 2001) Test 1 of Verbal Comprehension, age, semester half, fall or spring semester, gender, and race with self-derivation performance. The criterion for a multiple regression model were assessed (as discussed below).

Description of predictors and self-derivation performance in each individual study. Descriptive statistics of verbal comprehension scores, age, and self-derivation in each study are reported in Table 1. Descriptive statistics of race, gender, semester half, and semester season with self-derivation are reported in Table 2. Substantial variability was observed in self-derivation scores, with each study ranging from near 0% to near 100%.

Association between predictors and self-derivation performance. The primary aim of this study was to determine which study variables and individual participant measures are associated with self-derivation and the variability in scores. To examine the collinearity among the variables, I examined the variance inflation factor (VIF). The VIF was within an acceptable range, since no values were above five (Craney & Surles, 2002). The assumptions for a linear model were assessed using diagnostic plots in R that examined Cook's distance, residuals, scale

location plots, and Q-Q plots for linearity. No values were extreme (no residual greater than 3) and the Q-Q plot displayed a linear pattern.

With all assumptions met for a multiple regression, the model was constructed. The initial model included all the predictors. Verbal comprehension scores, $t(328) = 8.865$, $p < 0.001$, and age, $t(328) = 2.258$, $p = 0.02$, were significant predictors of self-derivation performance. The model was adjusted to include only these significant predictors in addition to study, as seen in Table 3. With verbal comprehension scores and study held constant, as age increased by 1, there was a predicted 3% increase in self-derivation performance. However, this finding must be interpreted with caution because of the narrow age range of participants (18 to 24 years old). Additionally, with age and study held constant, as verbal comprehension increased by 1 point, there was a 2% increase in self derivation. As such, there is evidence that both age and verbal comprehension have a positive relation with self-derivation. This adjusted model explained 36% of the variance in self-derivation performance, $F(7,328) = 28.23$, $p < 0.001$. Study remained in the analysis despite not having all statistically significant levels because it is fundamental for an IPD meta-analysis to include a grouping variable for each study. As a result, it is important to consider that study does explain a portion of the variance. Additionally, I found that study 231, $t(328) = 3.31$, $p = 0.001$, and study 248, $t(328) = 5.00$, $p < 0.001$, had significantly higher performances on self-derivation tasks than the baseline study (Study 189).

Additionally, to ensure that age and verbal comprehension scores were not correlated with one another, I ran a correlation test between the two variables. Though the two variables had a correlation significantly different from zero, $r(334) = 0.13$, $p < 0.05$, the correlation of 0.13 is quite low. This finding would indicate that a relation between the age and verbal comprehension with self-derivation was not due to between-predictor correlation.

Discussion

In the present study, I used an IPD meta-analysis to conduct a multi-regression to examine individual difference measures that could be associated with self-derivation performance in adults. I examined verbal comprehension scores on the Woodcock Johnson III (Woodcock, McGrew & Mather, 2001) Test 1 of Verbal Comprehension, age, semester half, fall or spring semester, gender, and race with self-derivation performance. Participant age and verbal comprehension scores were associated with self-derivation performance and explained a significant portion of variability. Additionally, study was included in the final model to account for participants coming from various studies. I also found that study 231 and study 248 had significantly greater performance on self-derivation than the other studies. This difference may be better explained in study 2 by looking at study-level factors.

Varga et al. (2019) was the first study showing that verbal comprehension was a significant predictor of self-derivation in adults. This IPD meta-analysis builds off of the previous findings by demonstrating a robust pattern with verbal comprehension and self-derivation. Specifically, this study found a significant positive association between verbal comprehension and self-derivation across studies (50%). Additionally, verbal comprehension accounted for a large portion of the variability in performance. As a result of these additional findings, there even more support for the strength of the relationship between verbal, semantic knowledge and self-derivation.

Though verbal comprehension and age were found to be significant across studies, it is important to interpret the relation between age and self-derivation performance with caution. In the present research I examined the self-derivation performance of a relatively limited age range, 18 to 24 years old. Currently, we do not know whether this pattern is stable throughout

adulthood. As a result, it would be interesting for future studies to explore the stability of this pattern. For example, we can examine whether the trends continue into ages beyond college students by including participants in their mid-to-late 20s and 30s.

In the next study, I aimed to build on these findings by examining self-derivation at the study level using a meta-analysis. This study-level meta-analysis is necessary to explore factors that are linked to study design and do not vary at the individual level, such as modality, number of stem facts presented, and assessment method. Such results will allow us to further study factors that are associated with self-derivation, and perhaps formulate more ideas as to why some people perform better than others on the task.

Study 2

Methods

Search Strategy and Study Selection

Once again, studies were collected from the server of the Bauer Memory Lab at Emory University in Atlanta, Georgia. Additionally, the study protocol and hypotheses for this study were preregistered on AsPredicted.org. The criteria for inclusion included criteria one through four from study one. This choice allowed for a greater number of studies to be included and did not restrict those without Woodcock Johnson III collection. Using these criteria, 11 studies were selected for inclusion out of the 110 available studies. Of these 11 studies, one has been published (200), one has been accepted for publication (248), one is under review (206), and eight are as yet unpublished (181, 189, 221, 221BC, 231, 239, 251, 253).

Participants

Upon compiling the data from each study, 543 participants were included (M = 19.56 years, SD = 1.12 years, 326 females) from the 11 studies. After examining the demographics

reported by each study, the compiled list of participants from each study consisted of 0.2% American Indian or Alaskan Native, 19.7% Asian, 11.4% Black or African American, 58.7% White or Caucasian, 3.7% Multi-racial, and 6.3% did not report. Each of the studies included in this study had received IRB approval when they were conducted. Participants received course credit

Study Design and Outcomes

The same data compiling process occurred as Study 1. Study number and proportion of self-derivation correct were still included in this analysis (see description in Study 1). In addition to these two factors, the following variables were extracted for analysis:

Modality. Modality refers to how the stem facts were presented to participants for encoding. Sentence modality refers to stem facts being presented one sentence at a time on a PowerPoint. One-word modality refers to stem facts being presented one word at a time in a PowerPoint format. Graphic/narrative modality consisted of participants being presented one stem fact in an informative graphic and the other stem fact pair in a three-sentence written excerpt. Two images with audio refers to two simple images being presented on a slide with an audio recording stating the stem fact. Finally, passage refers to using short paragraph pairs about a topic rather than single sentence stem facts.

Assessment Method. Self-derivation was tested in either an open-ended format or forced-choice format. Open-ended assessment required participants to speak or write the answer to a question assessing self-derivation, whereas forced-choice assessment required participants to select an answer from a list of possible answers.

Number of Facts to Learn. Each study varied in the number of stem facts presented to participants to learn. Pairs of stem facts (e.g., “Kierkegaard was the founder of Existentialism”,

"The self is the main focus of Existentialism") can be integrated to form novel facts (e.g., "Kierkegaard developed a philosophy that emphasized the self"). The number of facts presented to participants ranged from 10 total facts to 120 total facts (M = 58.8, SD = 33.1).

Statistical Analysis

To assess the relations between the study-level variables and self-derivation performance, a traditional meta-analysis was conducted. The effect size for each study was the calculated mean difference between the mean integration score when presented with two stem-facts and the mean integration score when presented with only one stem-fact. Using standard self-deviation theory, the control condition of one-stem fact presentation was assumed to have an effect size of zero and a standard deviation of zero. These values are based on the assumption that in order for self-derivation to occur, two stem facts are needed. Without the two stem facts, participants would not have adequate information to form the novel fact. In a separate examination of the studies, participants were found to have answered some of the one-stem integration questions; however, this indicates that participants had outside knowledge about the stem topic or were successfully able to guess the answer. As such, in order to assess the true treatment effect two-stem presentation versus one-stem presentation with no outside knowledge, zero was used in the analysis.

Using the metafor (Viechtbauer, 2010) package in R version 4.0.2., the escalc function was used to calculate the effects size (mean difference) and the standard error. To see whether self-derivation was significantly different from zero in each of the studies included in the analysis and to see the initial measures of heterogeneity, an initial model was constructed with none of the predictors. A random-effects-model was conducted using the Sidik-Jonkman tau

estimator for between-study variance. The Hartung-Knapp method was also used because it provides a more conservative estimate (Harrer et al., 2019).

After the initial model was constructed, I assessed whether modality, assessment method, and number facts explained any of the variation found in the initial model. This process was done by conducting a meta-regression by inputting each of the predictors. To start, I used the maximum likelihood and Knapp-Harung methods and inputted all three predictors. Afterwards, each predictor was used to construct their own model.

Results

In Study 2, I investigated which study-level variables predicted self-derivation performance. To achieve this goal, I conducted a study-level meta-analysis using a meta-regression. The study-level variables addressed in this model are modality, assessment method, and number of facts presented.

Examining self-derivation across each study with no predictors. As displayed in Figure 1 which shows the results of the initial model with none of the predictors, the effect size of each study was significantly above zero. This trend illustrates that self-derivation occurred in each of the included studies. Additionally, the I^2 of this model was 91%, $Q(10) = 115.69$, $p < 0.001$, meaning that 91% of the observed variance across studies comes from real differences between the studies, and not chance alone.

Assessment of predictors with proportion of successful self-derivation. To assess whether modality, assessment method, and number of facts explained any of the variation found in the initial model, I conducted a meta-regression. To start, I used the maximum likelihood and Knapp-Harung methods and included all three predictors. With this entry, the test of moderators was found to be not statistically significant, $F(6, 4) = 3.99$, $p > .05$. Additionally, the I^2 of this

model was still 91%, which is equal to the heterogeneity of the original model. Following this meta-regression, I assessed each of the predictors in their own model. No predictor was found to be statistically significant when assessed individually: modality, $F(4,6) = 3.83$, $p = 0.07$; assessment method, $F(1,9) = 2.14$, $p = 0.18$; and number of facts, $F(1,9) = 2.79$, $p = 0.13$. Figure 2 below reflects that each predictor did not reach a threshold that would deem it as an important and meaningful predictor in the final model. As such, though there is variation across studies in the performance of participants on the self-derivation tasks, there is no evidence that modality, assessment method, and number of facts presented are associated with this variability.

Discussion

The present study extended on the findings of Study 1 by examining study-level variables that could be related to self-derivation performance. The initial meta-analysis revealed a substantial amount of variation found across studies that cannot be explained by chance alone ($I^2 = 91\%$). Additionally, I found that each study evidenced that adults successfully self-derived new knowledge. This finding is supported by the fact that each of the effect size confidence intervals did not consist of the value zero. To determine if modality, number of facts presented to participants, and assessment method were significantly associated with self-derivation and explained this variation, I conducted a meta-regression using all three predictors and each predictor on their own. I found that modality, number of facts, and assessment method are not statistically significantly associated with self-derivation across the studies. As a result, I was not able to explain which study-level factors, if any, contribute to the extensive variability observed across studies.

General Discussion

The goal of this research was to identify individual-level and study-level factors that are associated with self-derivation performance and explain the extensive variability in performance. I had hypothesized that verbal comprehension, semester and the half of the semester, modality, assessment method, and number of facts to learn would be associated with self-derivation and the variability of performance. After conducting each analysis, verbal comprehension was significantly associated with self-derivation performance and variability. In addition, age was found to be a significant predictor, contrary to the hypothesis that it was a controlling factor.

First, I conducted an IPD meta-analysis to study verbal comprehension and semester and semester half with age, race, and gender as controlling factors. Using a multiple regression with study included to account for the different studies, I found that verbal comprehension and age were significantly associated with self-derivation and explained a sizeable proportion of individual variability in performance. As mentioned previously, the relation between verbal comprehension and self-derivation was observed across the studies, further strengthening the findings of Varga et al. (2019). Additionally, age was unexpectedly identified to have a significant positive relationship with self-derivation. This is the first reported finding of self-derivation performance increasing with age in adulthood. Together, verbal comprehension scores and age accounted for 36% of the variation in self-derivation performance. This high percentage indicate that these factors explain a portion of the vast range in self-derivation performance.

After conducting the IPD meta-analysis, I conducted a study-level meta-analysis to examine modality, assessment method, and number of facts presented. First, an initial model was constructed with none of the predictors. This model illustrated that there was substantial variability across each study in self-derivation performance, with 91% of the variability not

explained by chance alone. To further explore this variability, each of the predictors were assessed individually and together in a meta-regression. After completing this analysis, modality, assessment method, and number of facts presented were found to be not statistically significant, meaning that the amount of variability they explained was not statistically significant.

Ultimately, the IPD meta-analysis and study-level meta-analysis provided greater insight into the variability in self-derivation performance. As expected in my hypotheses, gender and race did not have a significant association with self-derivation. This finding is consistent with other studies that have found so such association (e.g., Varga, Esposito & Bauer, 2019). However, my results did not support previous findings by Porfido and colleagues (2020) regarding student study participation at the end of terms. As a result, my findings suggest that participants from the first half of the semester and the second half of the semester did not differ in performance. Acknowledging the findings of Porfido and colleagues (2020), perhaps most of the students who participated in the studies during the latter half of the semester were not showing effects from burnout or procrastination on study participation course credit. While I did not identify a significant relation between semester and semester half, our findings do not discredit the previous research.

Additionally, I was quite surprised to discover that modality, assessment method, and number of facts present were not statistically significantly associated with self-derivation performance across studies. Perhaps this was the case as a result of the limited number of studies available for analysis. For example, only one out of the 11 studies used forced-choice integration questions as the primary form of assessment. I completed a power analysis of subgroup differences, which indicated adequate power to detect an effect size in self-derivation (above 90%). However, the lower number of studies with certain conditions made it a challenge to

detect predictor significance, if one truly exists. As a result, it is important not to rule out modality and assessment type as possible predictors. Additionally, even though a wide range existed in the number of facts participants were exposed to across studies, perhaps having more than 11 self-derivation values to analyze would allow for the true effect size of the number of facts to be detected, contingent that one truly exits. As a result, future studies should address modality, number of facts presented, and assessment method to confirm the findings of this meta-analysis.

The findings outlined above provide us with a greater understanding of self-derivation because we were able to examine patterns across multiple studies. This meta-analysis confirmed previous findings about the importance of verbal comprehension, and increased our confidence in the relationship between verbal, semantic knowledge and self-derivation. Additionally, we now have greater insight into factors across studies that are associated with the vast variability in scores. By understanding what is associated with this variability, we can try to figure out ways to reduce the range and optimize performance. By investigating these relations, we can identify strategies to optimize learning. For example, we can investigate how to increase verbal, semantic knowledge in school aged children and young adults. Since my study confirmed previous findings regarding verbal comprehension, we know it could be beneficial to self-derivation performance to increase this semantic knowledge. Perhaps schools could introduce more vocabulary-based lessons over the course of childhood education. However, further research examining how to optimize verbal, semantic knowledge would be needed before implementing community-wide changes.

Limitations

While this study does contribute important findings, it is critical to consider the limitations of this study. Generalizability of this study is limited by the fact that each study included consisted of college students at the same private university. As a result, my findings are only directly applicable to this population. Additionally, this study utilized only a small number of studies. This quality could have resulted in weak power, reducing my ability to detect significant relation between the predictors and self-derivation. However, this limitation was unavoidable because there are a limited number of studies about self-derivation available to analyze that fit the parameters of this study. This study did benefit by having access to unpublished data, in addition to published data, avoiding publication bias all together. This availability also allowed me to have even more studies to include in the analysis. However, until more studies are produced examining the standard self-derivation paradigm in adults, we may not have the power to detect significant relationships between study-level predictors and self-derivation in a meta-analysis.

The findings of this study also help guide future research directions. Specifically, it would be beneficial to conduct research examining modality, number of facts, and assessment methods in their own respective study. As stated before, while these predictors were not statistically significant in this study, it could have been a result of the number of studies with each level of the variables (e.g., there was only one study that utilized the "narrative" modality). As a result, it would be critical to examine each of these predictors in their own study to confirm whether there is an association with self-derivation. Additionally, researchers would benefit from studying the relation between age and self-derivation further. Specifically, it would be helpful to include participants in their mid-to-late 20s and 30s to further explore whether the age trend remains when a broader range of ages are included. Keeping in mind that research is often done

in the University setting, perhaps graduate students could be recruited to be participants. This study would allow us to understand whether self-derivation improves into adulthood and beyond childhood. It would also provide us greater insight into the development of self-derivation and whether it continues to improve or plateau with age. Currently, it has been found that both adults and children can successfully self-derive novel information; however, they do so under different conditions. Adults can integrate rather successfully even when not prompted to do so, whereas children are less successful and many times their integration is sufficient for forced-choice but not open-ended performance (Bauer et al., 2020). Acknowledging these findings, it would be interesting to see whether adults continue to optimize this process or whether self-derivation no longer continues to improve.

Ways to improve verbal, semantic knowledge should be studied further. This study and others before have further reinforced the idea that verbal comprehension is closely related to self-derivation performance. Now that we have greater confidence about this relation, we should further study how to improve verbal comprehensions skills and examine whether improvement in these skills is associated with higher self-derivation performances. Conducting this type of study is critical to implementing academic changes that would be aimed at increasing indirect learning skills such as derivation. By further studying verbal, semantic knowledge improvement, we can explore the intricacies of the relation with self-derivation.

Conclusion

With the present research, I was able to accomplish my goal of compiling past research to determine individual measures and study design factors that are associated with performance on self-derivation tasks. While many questions still remain as to why some individuals are more successful at deriving novel facts than others, this study helped to establish preliminary findings.

By optimizing the process of self-derivation and understanding individual difference, we can start to identify ways to make individuals better indirect learners. Learning indirectly is critical to forming an extensive knowledge base, and by understanding how this knowledge base forms, we are able to become better learners. Ultimately, this study has been able to introduce a framework for future works examining self-derivation by identifying critical factors involved in self-derivation.

References

APA dictionary of psychology. (2020). American Psychological Association. https://dictionary.apa.org/

Baghurst, T., & Kelley, B. C. (2014). An examination of stress in college students over the course of a semester. *Health Promotion Practice*, *15*(3), 438-447. https://doi.org/10.1177/1524839913510316

Bauer, P. J., Cronin-Golomb, L. M., Porter, B. M., Jaganjac, A., & Miller, H. E. (2020). Integration of memory content in adults and children: Developmental differences in task conditions and functional consequences. *Journal of Experimental Psychology: General*, 1-20. http://doi.org/10.1037/xge0000996

Bauer, P. J., & Jackson, F. L. (2015). Semantic elaboration: ERPs reveal rapid transition from novel to known. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 271-282. https://doi.org/10.1037/a0037405

Bauer, P. J., & Varga, N. L. (2017). Similarity and deviation in event segmentation and memory integration: Commentary on Richmond, Gold, & Zacks. *Journal of Applied Research in Memory and Cognition*, *6*(2), 124-128. http://doi.org/10.1016/j.jarmac.2017.01.006

Craney, T. A., & Surles, J. G. (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, *14*(3), 391-403. https://doi.org/10.1081/QEN-120001878

**Cronin-Golomb, L. & Bauer, P.J. (2021) *Pre-testing as a facilitator of memory integration across representational formats*. manuscript in preparation. 253.

*Dugan, J. A., & Bauer, P. J. (2020). This should help with that: A behavioral investigation into self-derivation of knowledge about prescription medications. 206.

**Dugan, J. A. & Bauer, P. J. (2020) *Metacognitive judgments improve adults' self-derivation through memory integration*. manuscript in preparation. 231.

Ericson, J. M., Kravitz, D. J., & Mitroff, S. R. (2017). Visual search: You are who you are (+ a learning curve). *Perception*, *46*(12), 1434-1441. https://doi.org/10.1177/03010006617721091

Esposito, A. G., & Bauer, P. J. (2018). Building a knowledge base: Predicting self-derivation through integration in 6- to 10-year-olds. *Journal of Experimental Child Psychology*, *176*, 55-72. http://doi.org/10.1016/j.jecp.2018.07.011

Glass, A. L., & Sinha, N. (2013). Multiple-choice questioning is an efficient instructional methodology that may be widely implemented in academic courses to improve exam performance. *Current Directions in Psychological Science*, *22*(6), 471-477. https://doi.org/10.1177/0963721413495870

Grady, C. L., McIntosh, A. R., Rajah, M. N., & Craik, F. I. M. (1998). Neural correlates of the episodic encoding of pictures and words. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(5), 2703-2708. https://doi.org/10.1073%2Fpnas.95.5.2703

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2019). *Doing meta-analysis in R: A hands-on guide*. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/

Kelleher, M., & Cro, S. (2019). Individual participant data meta-analysis. In *Annual Cochrane Skin Group Meeting & CS-COUSIN*. Imperial College London.

Maricutoiu, L. P., & Sulea, C. (2019). Evolution of self-efficacy, student engagement and student burnout during a semester. A multilevel structural equation modeling approach. *Learning and Individual Differences*, *76*. https://doi.org/10.1016/j.lindif.2019.101785

Matthews, G., Sparkes, T. J., & Bygrave, H. M. (1996). Attentional overload, stress, and simulated driving performance. *Human Performance*, *9*(1), 77-101. https://doi.org/10.1207/s15327043hup0901_5

Porfido, C. L., Cox, P. H., Adamo, S. H., & Mitroff, S. R. (2020). Recruiting from the shallow end of the pool: Differences in cognitive and compliance measures for subject pool participants based on enrollment time across an academic term. *Visual Cognition*, *28*(1), 1-9. https://doi.org/10.1080/13506285.2019.1702602

Riley, R. D., Legha, A., Jackson, D., Morris, T. P., Ensor, J., Snell, K. I., White, I. R., & Burke, D. L. (2020). One-stage individual participant data meta-analysis models for continuous and binary outcomes: Comparison of treatment coding options and estimation methods. *Statistics in Medicine*, *39*(19), 2536-2555.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*(2), 378-389. https://doi.org/10.1016/j.neuron.2008.09.023

Varga, N. L., & Bauer, P. J. (2017). Young adults self-derive and retain new factual knowledge through memory integration. *Memory & Cognition*, *45*(6), 1014-1027.

**Varga, N. L., & Bauer, P.J. (2016). *Knowledge extension through memory integration: Factors and conditions that promote the long-term accumulation of knowledge*. manuscript in preparation.181 & 189.

**Varga, N.L., Cronin-Golomb, L., Doyle, K., Elder, J., Bauer, P.J. (2021) *General and specific prior knowledge support learning and integration of semantically-related information.* manuscript in preparation. 221BC.

Varga, N. L., Esposito, A. G., & Bauer, P. J. (2019). Cognitive correlates of memory integration across development: Explaining variability in an educationally relevant phenomenon. *Journal of Experimental Psychology: General*, *148*(4), 739-762.

*Wilson, J. T., & Bauer, P. J. (2020). Prompt-facilitated learning: The development of unprompted memory integration and subsequent self-derivation. 248.

Wilson, J. T., & Bauer, P. J. (2021). *Comparing retrieval practice and self-derivation: Implications for retention and extension of information.* manuscript in preparation. 251.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III test of cognitive abilities*. Riverside Publishing.

Writing a literature review. (2020). Purdue Online Writing Lab. https://owl.purdue.edu/owl/research_and_citation/conducting_research/writing_a_literature_review.html

**Yi, M., Dugan, J. A., & Bauer, P.J. (2018). *The role of metacognition in self-derivation of new factual knowledge across episodes of learning through infographics and text*. manuscript in preparation. 239.

Note: * denotes studies that are accepted for publication or under review, ** denotes studies that are unpublished.

Table 1:

Study 1 Descriptive Statistics for Quantitative Measures by Study

		Verbal Comprehension		Age		Self-Derivation Score		
Study	n	M	SD	M	SD	M	SD	Range
Study 189	23	57.65	3.97	19.63	1.26	0.45	0.25	0.07-1.00
Study 200	114	57.92	4.17	19.76	1.15	0.51	0.21	0.03-0.93
Study 206	53	46.04	5.58	19.89	1.24	0.25	0.26	0.00-1.00
Study 231	100	47.73	4.82	19.26	0.94	0.39	0.21	0.00-0.91
Study 248	28	54.43	4.65	19.03	0.86	0.65	0.24	0.13-1.00
Study 253	22	55.82	3.57	19.31	0.75	0.29	0.21	0.00-0.60

Table 2:

Study 1 Descriptive Statistics for Categorical Measures

			Self-Derivation Score		
Variable		N	Mean	SD	Range
Race	American Indian or Alaskan Native	1	0.27	NA	NA
	Asian	89	0.41	0.24	0.00-1.00
	Black or African American	41	0.41	0.27	0.00-0.88
	White/Caucasion	190	0.44	0.25	0.00-1.00
	Multi Race	16	0.35	0.29	0.00-0.93
	Did Not Disclose	3	0.58	0.20	0.36-0.76
Gender	Female	222	0.42	0.25	0.00-1.00
	Male	118	0.44	0.25	0.00-1.00
Semester	First Half	214	0.42	0.26	0.00-1.00
	Second Half	119	0.43	0.24	0.00-1.00
	NA	7	0.52	0.18	0.24-0.76
Semester Season	Fall	168	0.44	0.25	0.00-1.00
	Spring	165	0.40	0.26	0.00-1.00
	Summer	7	0.40	0.18	0.24-0.76

Table 3

Study 1 Summary of Multiple Regression Analysis

Variable	Level	R^2	β	VIF	p
Study				2.42	
	200		0.05		0.27
	206		0.05		0.37
	231		0.17**		0.001
	248		0.29**		<0.001
	253		-0.11		0.08
Age			0.03*	1.09	0.01
Verbal Comprehension			0.02**	2.29	<0.001
Total Adjusted R^2		36%			

Note: VIF = variance inflation factor. * p <0.05, **p < 0.01.

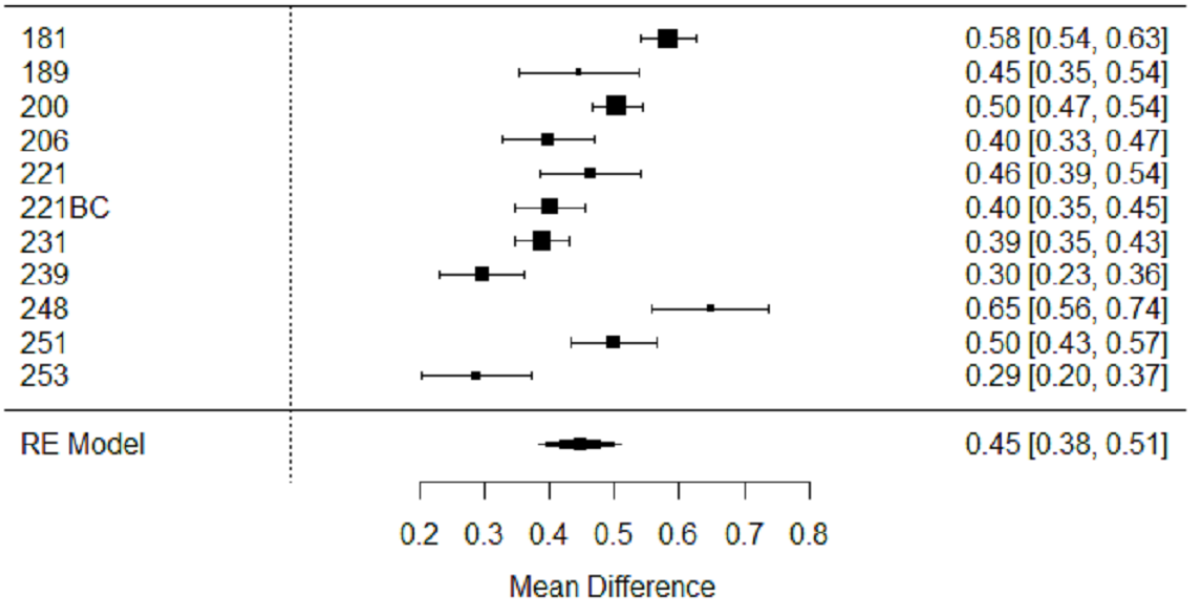

Figure 1. Forest plot depicting the effect sizes of each study and predicted effect size. The mean difference and 95% confidence interval are listed for each study and the overall model. The server number for each study is included as the study identifier. Citations can be found in the references.

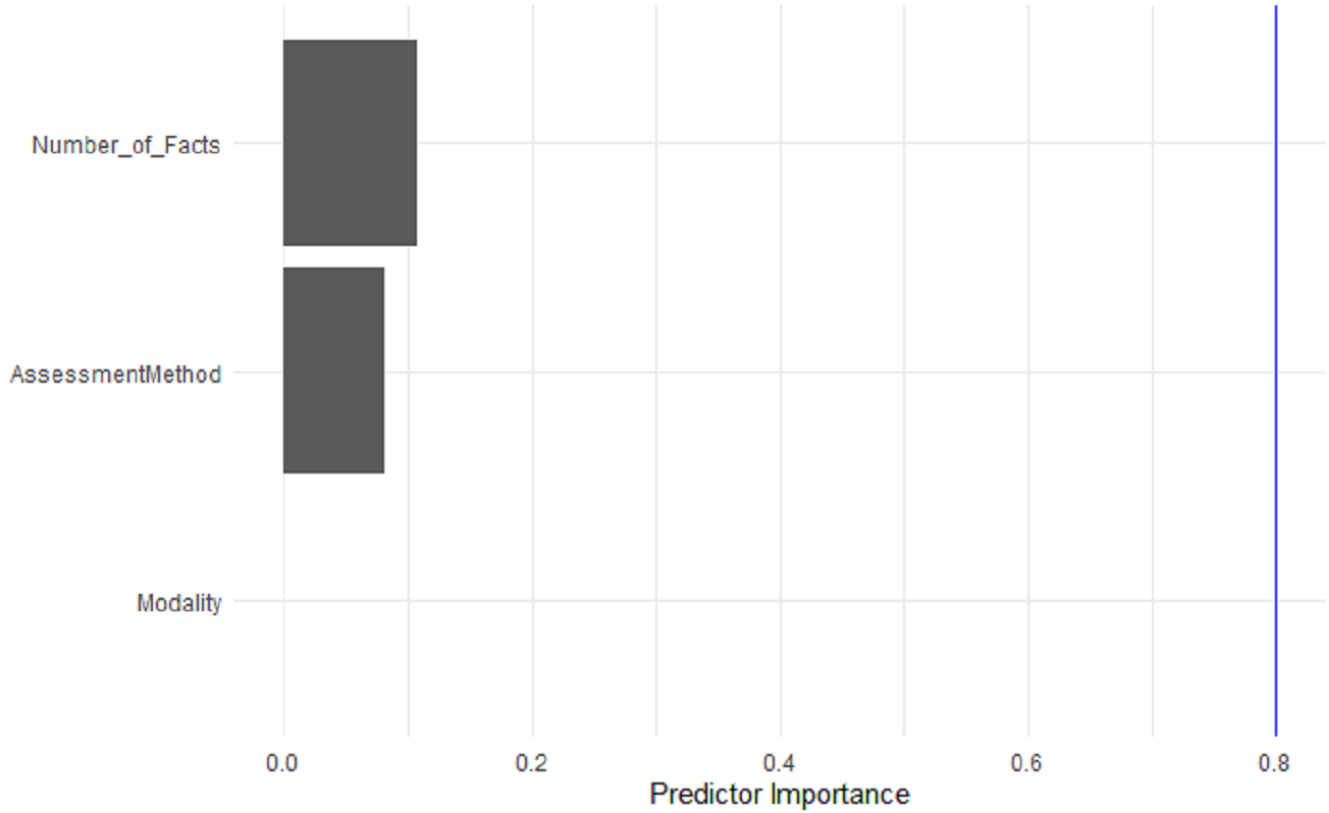

Figure 2. Predictor importance plot depicting the ranking of predictors from the dmetar package (Harrer et al., 2019).