Studies in Second Language Acquisition (2025), **47**, 1148–1169
doi:10.1017/S0272263125100946

METHODS FORUM

Interpreting the semi-partial correlation as a multiple regression-bound (not a bivariate) metric: A methods-oriented response to Papi and Teimouri's (2024) response to Al-Hoorie et al. (2024)

Joseph P. Vitta[1], Paul Leeming[2] and Christopher Nicklin[3]

[1]Waseda University - Global Education Center, Tokyo, Japan; [2]Kindai University, Osaka, Japan and [3]The University of Tokyo, Center for Global Education, Tokyo, Japan
Corresponding author: Joseph P. Vitta; Email: vittajp@waseda.jp

(Received 06 January 2025; Revised 14 May 2025; Accepted 25 May 2025)

Abstract
Al-Hoorie, Hiver, and In'nami (2024) challenged the validity and corresponding validation processes of L2 Motivational Self System (L2MSS) research. A component of this challenge included claims of weak discriminant validity due to high correlations among L2MSS constructs. Papi and Teimouri (2024) countered by using semi-partial correlations to control for other L2MSS constructs, finding weak-to-moderate associations, which they claimed mollified potential discriminant validity concerns. In this methods-oriented response paper, we present a historical case that semi-partial correlations should be viewed within the context of multiple regression analysis, not as a standalone bivariate metric. Challenging Papi and Teimouri's approach, we suggest that their method does not adequately address discriminant validity issues. Furthermore, when their semi-partial correlations are treated as multiple regression models, Al-Hoorie et al.'s concerns remain valid. Finally, we demonstrate that L2MSS literature does not support the assignment of outcome and predictor variables in Papi and Teimouri's semi-partial correlations when correctly considered as multiple regression models.

Keywords: Multiple Regression; Validation; Semi-partial Correlation; L2 Methods Reform; L2MSS

Introduction

The debate around Al-Hoorie et al.'s (2024) challenge to L2 Motivational Self System (L2MSS: Dörnyei, 2009) theory has been robust, stimulating, and fruitful for the field. From this ongoing discussion, opportunities have emerged to inform and extend the field's collective expertise. Henry and Liu (2024), for example, when responding to

Al-Hoorie and colleagues, challenged the field to do more to understand the seminal literature and concepts therein when considering how to operationalize measurements pertaining to what motivates language learners. This debate has also led to calls for the enhancement of our methodological approaches to investigating language learners' motivation. To wit, Al-Hoorie et al.'s challenge was underpinned by using exploratory factor analysis in addition to confirmatory factor analysis to support their claims of validity concerns regarding the L2MSS along several lines. Such pairing of the procedures with different samples for each constitutes a gold standard cross-validation approach that has strong support in the social science psychometric literature (Henson & Roberts, 2006; Worthington & Whittaker, 2006) and has been used recently in psychology of language learning (PLL) research (e.g., Leeming, Vitta, Hiver, Hicks, McLean, & Nicklin, 2024). To cite another example, in responding to Papi and Teimouri (2024), Oga-Baldwin (2024) rightly highlighted the advantages that latent variable structural equation modeling (SEM) offers over its mean-driven ordinary least squares (OLS) regression alternative. It bears noting, however, that mean-driven OLS regression is a type of SEM pathway analysis where the latent variables are constructed via aggregation of indicators (Kline, 2023). Both Al-Hoorie et al. (2024) and Papi and Teimouri (2024), furthermore, appeared to use such aggregates of Likert scales in OLS modeling in their reports. In sum, the debate started by Al-Hoorie et al. has fostered several strands of discussion that serve to enhance the field's practices and knowledge.

In this response to Al-Hoorie et al. (2024) and Papi and Teimouri (2024), we demonstrate that semi-partial correlations are not suitable for establishing discriminant validity given their approximately 100-year-old use as a metric of multiple regression (Dunlap & Cureton, 1930; Wright, 1921)[1]. This demonstration involves a reconsideration of Papi and Teimouri's semi-partial correlation analyses involving Al-Hoorie et al.'s data. In such models, the semi-partial correlation metric describes a predictor variable's (*X1*) association with the outcome variable (*Y*) after removing the association that other predictors (*X2, X3…*) have on *Y* (see Aloe, 2015)[2]. Our response begins by demonstrating the historical case of viewing the semi-partial correlation as a multiple regression-bound metric and by doing so, demonstrating how the metric cannot describe the bivariate associative relationship between two measures (Aloe, 2014; 2015). Next, the semi-partial correlations presented in Papi and Teimouri (2024) are reconsidered as multiple regression models, which in reality they are. This reconsideration reviews L2MSS theory, where there are weak cases at best for the observed ascription of outcome (dependent) and predictor (independent) variable labels in said models. This paper also presents Papi and Teimouri's (2024) semi-partial correlations using the underlying logic of the metric. When conducting such an analysis, it is essential to consider both the unique predicted variance in the multiple regression model attributable to each predictor and the predicted variance in the model that is ascribable to the common variance among the predictors (Abdi, 2007; Aloe, 2014; 2015). These analyses demonstrate that Al-Hoorie et al.'s

[1]Multiple regression is a multivariable (multiple predictors/independent variables) modeling approach underpinned by the general linear model, which drives all parametric testing (Cohen et al., 2003; Plonsky & Oswald, 2017). Multivariate denotes models with multiple outcomes or dependent variables. This report focuses almost exclusively on multiple regression, given Papi and Teimouri's (2024) use of semi-partial correlations with one apparent outcome variable.

[2]As noted by Aloe (2015), it is also possible to think of semi-partial correlations in a manner where the metric adjusts for the effects of the other predictors on the predictor in question but not *Y*.

discriminant validity concerns hold, as the common variance among predictors still accounts for an overwhelming amount of the explained outcome variable variance in the models.

At the outset, it is important to note that this response is undertaken to contribute to an ongoing and productive discussion that has the potential to benefit future language learning motivation research in particular and SLA research in general. It is not our intent to criticize, and we acknowledge that no work is perfect, including the reports comprising this debate. Furthermore, this response serves as a purely methods-focused response to Papi and Teimouri's (2024) use of semi-partial correlations (and thereby multiple regression). It is noteworthy that their response to Al-Hoorie and colleagues (2024) covered a range of perceived issues, including the use of exploratory factor analysis to test discriminant validity and their operationalization of L2MSS constructs. Although we acknowledge that these issues are currently being debated, this response focuses exclusively on Papi and Teimouri's use of semi-partial correlations.

The inseparable relationship between semi-partial correlations and multiple regression

Papi and Teimouri (2024) used semi-partial correlations to describe the association between several sets of two L2MSS variables in an effort to show that their associations were not as strong as Al-Hoorie et al. (2024) claimed. Consider the following excerpt from their report:

> To statistically examine the discriminant validity of Ideal L2 Self and Linguistic Self-Confidence…we ran a semipartial correlation…(see Lawson & Robins, 2021) while controlling for the correlation between Vividness of Imagery (which only represents the vision aspect of their Ideal L2 Self) and Linguistic Self-Confidence…the analysis showed a modest correlation of (r = .28) between L2 Self-Confidence and Ideal L2 Self, representing nearly 8% of the shared variance and leaving 92% of the variance unexplained. This clearly rejects the argument that "response to the Ideal L2 Self might be driven by a belief in ability rather than an actual–ideal discrepancy" (p. 4)

Their salient claims seem to be (a) controlling for Vividness of Imagery, the association between Ideal L2 Self and Linguistic Self-Confidence is small enough to assuage concerns about discriminant validity; (b) Lawson and Robins (2021) offer theoretical (or methodological) support for their use of semi-partial correlations; and (c) the semi-partial correlations can be used to home in on two variable's association ("92% of the variance unexplained"). When reviewing the historical precedent to view semi-partial correlations as a multiple regression-bound metric, however, Papi and Teimouri's use of semi-partial correlations to make these claims appears to be problematic along three lines. First, multiple regressions and semi-partial correlations are not "similar" to each other (Papi & Teimouri, 2024, p. 4), but are a part of a singular modeling approach to inferential testing. In this sense, semi-partial correlations can be considered a multiple regression-bound metric and, in turn, require strong theoretical support for the assignment of outcome and predictor variable status (e.g., Englehart, 1936; Aloe, 2014; 2015). Second, semi-partial correlations do not describe the bivariate relationship between two variables in isolation, and they must be considered within the broader multiple regression model as one of three multiple regression predictor-level effect

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

metrics. Finally, the logic underlying semi-partial correlations, where explained or predicted variance in the model is either unique to a given predictor or common (or shared) among predictor(s) in the model (Nathans, Oswald, & Nimon, 2012), has implications for interpreting their use in the manner Papi and Teimouri propose.

Semi-partial correlations as a multiple regression-bound metric and the importance of theory in multiple regression modeling

Papi and Teimouri (2024) stated that the relationship between semi-partial correlations is "similar to multiple regression analysis" (p. 4) and cited Lawson and Robins (2021) to support their use of the metric. Papi and Teimouri did appear to accurately summarize Lawson and Robins (2021) position that when considering highly related "sibling" constructs' associations with each other and/or other variables of interest, "researchers can calculate a part correlation (also called semi-partial correlation; Abdi, 2007), which involves holding the sibling construct constant only for the focal variable (not the outcome) as in a multiple regression" (p. 359). In other words, Lawson and Robins (2021) seemed to propose, with caution (see their hedging on this matter on p. 359) citing Abdi (2007), that semi-partial correlations were appropriate for the manner in which Papi and Teimouri used the metric. It is noteworthy, however, that Lawson and Robins (2021) only mentioned semi-partial correlations once in their paper's main text.

A review of Abdi (2007), however, makes it clear that Lawson and Robins might have misunderstood Abdi's position. Essentially, semi-partial correlations are a multiple regression-bound metric as Abdi clearly stated: "The semi-partial regression (correlation) coefficient—also called part correlation—is used to express the specific portion of variance explained by a given independent variable in a multiple linear regression analysis (MLR)" (p. 1). To state it simply, semi-partial correlations are not similar to multiple regression but instead are a component of the model's computations and reporting. Furthermore, at no point did Abdi raise the notion of focal measurement or sibling construct measurement. Aloe (2014; 2015) did employ the term "focal," but in the context of semi-partial correlations where the focal predictor (i.e., the predictor for which a semi-partial correlation was being computed) was juxtaposed to other predictors in the model. In neither Abdi (2007) nor Aloe (2014; 2015) did the use of semi-partial correlations to address discriminant validity appear, which corresponds to the concerns raised by Oga-Baldwin (2024), who argued for the use of computations found within "stringent factor analyses" (p. 7) to assess discriminant validity. Therefore, when Papi and Teimouri stated that "the semi-partial correlation between the focal and outcome variable will be shared variance independent of the correlation between the sibling constructs, supporting the incremental validity of the focal variable" (p. 4), they were accurately summarizing Lawson and Robins (2021). However, they also potentially misconceptualized multiple regression, to which semi-partial correlations are bound, as a scale/measurement validation tool instead of an inferential hypothesis testing procedure, where outcome and predictor variable status have a clear theoretical defense.

The historical case, extending from Abdi (2007) and Aloe (2014; 2015), to view the semi-partial correlation as a multiple regression metric, is easily made in the literature. Cohen, Cohen, West, and Aiken (2003), for instance, presented semi-partial correlations as a means to understand each predictor's unique contribution to a multivariable/multivariate model's omnibus effect size, R^2 (see Section 3.3.2, p. 72). Furthermore, the historical precedent for this position, along with a clear rationalization for the assignment of outcome and predictor variable structure, dates back a century. Englehart (1936), citing Wright (1921), for example, reviewed the metric among other multiple regression metrics, such as the standardized regression coefficient, in the context of models featuring a dependent variable (i.e., the effect) and independent variables (i.e., the causes or predictors of the dependent variable).

To cite another historical example, Dunlap and Cureton (1930) reviewed the semi-partial correlation alongside other proposed multiple regression-bound metrics such as (multiple) standardized regression coefficients and partial correlations. In this example, the outcome variable was children's intelligence, and the two predictors were the parents' intelligence and the home environment. As with Aloe (2014; 2015), Dunlap and Cureton demonstrated how semi-partial correlations isolate the component of each predictor that is uncorrelated with other predictors and then capture the association of these uncorrelated components with the outcome to address a predetermined question of causation. From the Englehart (1936) and Dunlap and Cureton (1930) examples comes the presumption that multiple regression has theoretical support for the assignment of outcome and predictor variables. Both use terms such as "cause and effect" and "causation." It is important to note that the underlying math of multiple regression does not change if correlational hypotheses, as opposed to experimental ones, are tested, but the assumption that theory supports the assignment of Y and X variables still holds. A useful example of this in the SLA literature is lexical sophistication research. Kyle and Crossley (2015; 2016) and Eguchi and Kyle (2020), for instance, predicted productive language proficiency scores (Y) with different metrics of lexical sophistication, including frequency ($X1$) and age of acquisition ($X2$), using multiple regression modeling. In these studies, the literature reviews established theoretical arguments for why the proposed lexical sophistication metrics would account for variation in the outcome variable(s).

In sum, the semi-partial correlation is a metric that is bound to multiple regression analysis, which is an inferential hypothesis testing process. Multiple regression is multivariable in nature and has strong theoretical support for the proposed structure of outcome and predictor variables. To understand semi-partial correlations' logic and how the metric works in the context of multiple regression, Abdi (2007) is now reviewed in a demonstrative manner.

The semi-partial correlation as one of three multiple regression predictor-level (partial) effect metrics

As noted by Aloe (2014; 2015), multiple regression modeling results in three metrics that can be used to describe each predictor's contribution to the model: (i) the standardized regression coefficient (b^*), (ii) the partial correlation, (iii) the semi-partial correlation (for historical support of this position; see Dunlap & Cureton, 1930; Englehart, 1936; McNemar, 1962; Velicer, 1978). These metrics, along with the omnibus effect size for multiple regression, R^2, can be calculated using the bivariate correlations among variables in the model (e.g., Dunlap & Cureton, 1930; Aloe, 2014). Demonstrative data (see Table 1) from Abdi (2007), who discussed all three metrics, is used to illustrate this process:

From these data, it is possible to calculate the correlations among the three variables (left unrounded to confirm Abdi's reporting): $r(Y.X1) = .8025$; $r(Y.X2) = .9890$; $r(X1.X2) = .7500$. The first two correlations are called *zero-order correlations* and these represent the bivariate association each predictor has with the outcome variable. The third correlation is the association between the two predictors. If the predictors are

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

Table 1. Demonstration data from Abdi (2007).

Case	*Y*: Memory span	*X1*: Age	*X2*: Speech rate
1	14	4	1
2	23	4	2
3	30	7	2
4	50	7	4
5	39	10	3
6	67	10	6

Note: Abdi (2007) assigns *X2* a label of T later in the chapter, but *X2* is retained here.

uncorrelated (i.e., $r[X1.X2] = .00$), the zero-order correlation value will equate to the values for the three multiple regression predictor-level (sometimes called "partial" [Aloe, 2015]) metrics presented above (Aloe, 2014; Velicer, 1978). In Abdi's case, the predictors are correlated, and this value is used to "correct" the zero-order value(s) to produce the three metrics commonly used to conceptualize each predictors' unique contribution to the multiple regression model (see the examples from the SPSS and JASP outputs for the Table 1 data in Figures 1a and 1b). First is the standardized regression coefficient (b^*). This metric predicts standard deviation (*SD*) movement in the outcome variable. In the example, for every +1*SD* movement in age, +14% *SD* of an

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	1.667	3.598		.463	.675	-9.782	13.116					
	X1_Age	1.000	.725	.140	1.379	.262	-1.307	3.307	.803	.623	.092	.438	2.286
	X2_Speech Rate	9.500	1.087	.884	8.736	.003	6.039	12.961	.989	.981	.585	.438	2.286

a. Dependent Variable: Y_Memory_Span

Figure 1a. SPSS multiple regression output using Abdi's (2007) demonstration data.

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
H_1	Regression	1822.000	2	911.000	110.054	0.002
	Residual	24.833	3	8.278		
	Total	1846.833	5			

Note. The intercept model is omitted, as no meaningful information can be shown.

Coefficients ▼

Model		Unstandardized	Standard Error	Standardized	t	p	95% CI Lower	95% CI Upper	Collinearity Statistics Tolerance	Collinearity Statistics VIF
H_0	(Intercept)	37.167	7.846		4.737	0.005	16.998	57.336		
H_1	(Intercept)	1.667	3.598		0.463	0.675	−9.782	13.116		
	X1_Age	1.000	0.725	0.140	1.379	0.262	−1.307	3.307	0.438	2.286
	X2_SpeechRate	9.500	1.087	0.884	8.736	0.003	6.039	12.961	0.438	2.286

Part And Partial Correlations

Model		Partial	Part
H_1	X1_Age	0.623	0.092
	X2_SpeechRate	0.981	0.585

Note. The intercept model is omitted, as no meaningful information can be shown.

Figure 1b. JASP multiple regression output using Abdi's (2007) demonstration data.
Note1: "Part" is another term for semi-partial correlations (Aloe, 2014; 2015).
Note2: Values equate to Abdi's reporting, but are slightly different due to rounding.
Note3: The SPSS and JASP outputs are provided to promote veracity of our claims and to provide further support confirming the convention of standardized beta (b^*), partial correlation, and semi-partial correlation as multiple-regression bound metrics.

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

upward movement in the outcome variable, memory span, can be attributed (or ascribed) to the predictor. Compare this relatively low value with the zero-order correlation of r = .8025. This is the process of the "math" behind multiple regression. Age's association with the memory span outcome variable has been reduced in the multiple regression model because it is highly correlated with a predictor (i.e., speech rate) that has a stronger bivariate association with said outcome variable.

Second is the partial correlation. This metric is interpreted in terms of the strength of association between a predictor variable and the outcome variable in the multiple regression model, fully controlling for other predictor variables and their associations with the outcome variable (Abdi, 2007). Because this value is based on a full adjustment for all variables, it is often less conservative than the semi-partial correlation (i.e., a larger effect reported with less nuance regarding the isolation of unique predictor effects [Dunlop & Cureton, 1930; Abdi, 2007]). What is implied by "less nuance" is that the full adjustment driving the partial correlation does not allow for the exploration of each predictor's unique contribution to the model to the same extent as the semi-partial correlation and the historical precedent for this viewpoint is almost a century old (see e.g., Dunlap & Cureton, 1930).

The semi-partial correlation, the third predictor-level multiple regression metric, differs from the partial correlation in that the adjustment is on the predictor side (hence "semi"), where common variance among the predictors is identified and then excluded, and then the uncorrelated or unique variance for each predictor has its association with the outcome variable calculated. As discussed by Dunlap and Cureton (1930) and summarized by Aloe (2014), this metric is useful in being the most conservative in relation to identifying what is truly unique to each predictor in relation to model predictiveness (see Figures 1a and 1b). This is because the semi-partial correlation values are lowest relative to beta and partial correlations (e.g., age: b^* = .14, partial correlation = .62, semi-partial correlation = .09). A caveat to this observation, however, is that the three partial effect metrics for multiple regression and other associated metrics such as Pratt (1987) product measure are derived from differing logical stances and mathematical assumptions (Aloe; 2014; Nimon & Oswald, 2013). Therefore, such cross-comparisons are to be made in a cautious manner. As a final note, zero-order, partial, and semi-partial correlations can be squared to conceptualize explained variance (Abdi, 2007), while convention usually only allows for b^* to be interpreted as a predictive *SD* movement inference.

Unpacking the mechanics and logic underlying semi-partial correlations

The different values for the b^*, partial correlation, and semi-partial correlation metrics for each predictor in Figures 1a and 1b touch on the observation that there are different ways to unpack partial or predictor-level effects in multiple regression modeling (and multivariable/multivariate modeling in general). This point was made by both Englehart (1936) and Dunlap and Cureton (1930), with the latter stating that other semi-partial correlation calculations "could be devised to meet other needs" (p. 679). As Papi and Teimouri (2024) used semi-partial correlations in their report, we engage with the formula and underlying logic of the metric in the context of multiple regression modeling using Abdi's (2007) example. It should be noted, however, that different approaches exist. Nimon and Oswald (2013), for instance, presented several different ways to understand or decompose the omnibus effect of multiple regression in relation to each predictor's contribution. Approaches such as Pratt's (1987) product measure

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

exist, where each predictor's b^* and zero-order correlation values are multiplied and the aggregate of these values will always equal R^2 (see use in Vitta, Nicklin, & Albright, 2023). This approach is less conservative and more straightforward than the logic of semi-partial correlations because the formula of b^* is used to avoid the question of common variance among predictors and to parse all R^2 variance out among each predictor. The Pratt product measure was challenged by Nimon and Oswald for situations where the model has too many variables and suppression renders Pratt values almost meaningless. The overarching point is that we do not present the semi-partial correlation as the sole approach to understanding the results of multiple regression, and this demonstration is to facilitate a robust response to Papi and Teimouri (2024).

Returning to Abdi's (2007) demonstrative data, one can calculate the semi-partial correlation for *X1* (age) using the following formula (see Aloe 2014; 2015 for confirmation of the formula):

$$\text{Semi-partial correlation } (Y.X1) = ((r[Y.X1] - (r[Y.X2]*r[X1.X2]))/((1 - r^2[X1.X2])^{.5}))$$

Inserting the values from Abdi (2007), one can calculate and see that the resulting values match Figures 1a and 1b:

$$\text{Semi-partial correlation}(Y.X1) = ((.8025 - (.9890*.7500))/((1 - .7500^2)^{.5})) = .092$$

As a correlation value, this can be squared to calculate that .85% of the outcome variable's predicted variance is unique to age. Speech rate can be processed similarly (it takes the place of *X1* in the formula above; age is *X2*) to calculate that it uniquely predicts 34.22%. These calculations, however, are not the end of the proverbial story

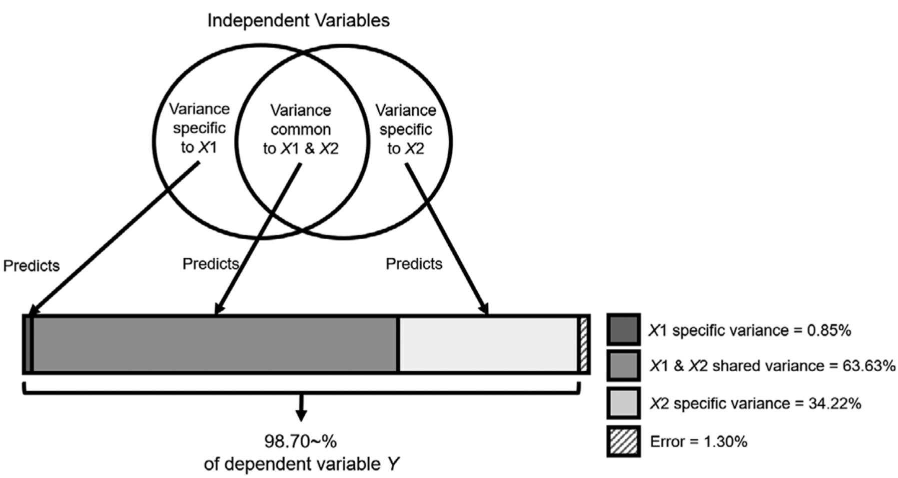

Figure 2. Visualization of the interpretation of the semi-partial correlations and omnibus multiple regression from Abdi's (2007) data.
Note: Adapted from Abdi (2007, p. 8) but values reported here are slightly different due to JASP/SPSS rounding.

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

when using semi-partial correlations. The multiple regression has an omnibus effect of 98.70% (R^2). As Figure 2 illustrates (recreated from Abdi, 2007), the variance not uniquely predicted by X1 and X2 is predicted by the common variance among the two predictors. Abdi demonstrated that this extrapolation is somewhat straightforward. By subtracting the squared values of all predictors' semi-partial correlations from the omnibus R^2 value for a given regression model, one can calculate the predicted variance in the model that is accounted for by the common variance among the two predictors. In a technical sense, the semi-partial correlation for a predictor is really the correlation between its residual, when all the other predictors in the model account for it as an outcome variable (see Supplementary Material A), and the actual outcome variable *in the multiple regression* (see McNemar, 1962). As the residual represents the predictor's independent variance relative to the other predictors, the claims of "unique" and "common" variance using the logic supporting semi-partial correlations in relation to multiple regression become clear.

The above point is what Papi and Teimouri (2024) might have missed in their use of semi-partial correlations. One cannot say that the association between age and memory span only accounts for .85% of the explained variance in the model, with 99.15% of it unexplained. Instead, age predicts 64.48% (.85% + 63.63%) of the memory span's variance in the model, either uniquely or in tandem with speech rate. The semi-partial correlation value does allow us to question the usefulness of age in the model, and it does help in considerations of causality (Dunlop & Cureton, 1930). The metric cannot be used, however, to argue that age and memory span are unrelated. The notion of "controlling for *X2*" using the logic of semi-partial correlations does not mean that we have "deleted" the variance of the controlling variable (i.e., *X2*). Instead, we have conservatively parsed out common and unique variance with both common and unique predictor variance contributing to the model's omnibus effect as illustrated in Figure 2.

Reconsideration of Papi and Teimouri's (2024) semi-partial correlations

The four semi-partial correlation sets presented by Papi and Teimouri are reconsidered as four multiple regression models. For each model, the proposed structure (i.e., assignment of outcome [dependent] and predictor [independent] variable status) is considered along with how variance is accounted for in the model. Because Papi and Teimouri used semi-partial correlations in their response, this consideration of variance allocation is presented using the underlying logic of the metric presented in the preceding text (see reviews in Abdi, 2007; Aloe, 2014), where the variance common among predictors in the model contributes to the omnibus effect (i.e., R^2) along with the unique contribution of each predictor. These analyses were conducted using data shared by Al-Hoorie et al. (2024) on the Open Science Framework (OSF: https://osf.io/7c8qs/), and we were able to recreate the effects of interest reported in both Al-Hoorie et al.'s and Papi and Teimouri's studies. In the interests of open science and transparency, we have also shared our aggregated scales of the original data on the Open Science Framework (OSF: https://osf.io/vjf6t/). It bears repeating, however, that there are different approaches to decomposing omnibus effects in multiple regression and other multivariable/multivariate models (Nathans et al., 2012; Nimon & Oswald, 2013).

To frame this reconsideration of the four models, it is important to first review L2MSS theory and consider where its constructs have consistently been proposed as predictors or antecedents of tangible L2 outcomes (e.g., Dornyei & Chan, 2013; You, Dörnyei, & Csizèr, 2016). Although it is important to consider the relationship between different parts of the L2MSS and how different constructs may relate to each other, it must be remembered that Dörnyei (2009) devised the L2MSS model to understand how motivation would influence language learning outcomes, including the effort expended on learning a language, which is often now analyzed under the guise of engagement (e.g., Hiver, Al-Hoorie, Vitta, & Wu, 2024), and also gains in proficiency.

Rather than breaking down the model to see how certain constructs may predict others (e.g., self-confidence predicting the Ideal L2 Self), the key concern should be how the variables connect in the model of the L2MSS to ultimately predict criterion (outcome) variables. Examples of this can be found in Dörynei's research. For example, Dörnyei and Chan (2013) were interested in the relationship between Imagery Capacity and Ideal L2 Self, but it is important to note that the aim of the study was to ascertain how these things impacted outcomes in the classroom, which in this case were intended effort and course grades. You et al. (2016) conducted a large-scale study investigating the L2MSS in China, and again, the outcome was the intended effort of the students to learn English. Although mono-method bias is an issue in a number of these studies, it is clear that the primary focus of the research is how the L2MSS can be used to predict students' learning behavior, rather than what seems to have become the focus in some research, which is on how parts of the model relate to each other with little or no consideration to the outcome of motivated language learning. While Papi and Teimouri (2024) discussed sibling constructs, it should also be noted that, to the best of our knowledge, Dörnyei never used this term, and many of the claims regarding sibling and parental constructs are speculative.

Therefore, the use of semi-partial correlations and, thus, multiple regression is questionable because such modeling presumes clear theoretical arguments for the ascribing of outcome and predictor variable status. As a final note, we do not imply that multiple regression and other inferential tests should never be used in an exploratory manner, but such use should be explicitly labeled as such to avoid the questionable research practice(s) of forming theory from already known results (Larsson, Plonsky, Sterling, Kytö, Yaw, & Wood, 2024). Papi and Teimouri (2024) appeared not to have made any such exploratory claims in their use of semi-partial correlations, and even if they had, it is difficult to imagine how exploratory modeling with multivariable inferential tests could be employed to challenge Al-Hoorie et al.'s (2024) claims about the L2MSS.

Model 1: Ideal L2 Self (*Y*), Linguistic Self-Confidence (*X1*), and Vividness of Imagery (*X2*)

The first semi-partial correlation that Papi and Teimouri (2024) presented was that between Ideal L2 Self (*Y*: outcome variable) and Linguistic Self-Confidence (*X1*: predictor variable 1), controlling for Vividness of Imagery (*X2*: predictor variable 2). While the report did not specify outcome and predictor variable assignment, one can conclude that Ideal L2 self was intended as the outcome when they stated "Dörnyei

Table 2. Abridged reporting of Model 1's (Y: Ideal L2 Self) multiple regression.

Predictor	b^*	r_{sp}	VIF
X1. Linguistic Self-Confidence	.39	.28	2.00
X2. Vividness of Imagery	.58	.41	2.00

Note 1. Observed Durbin-Watson value (1.94) and residual visualizations suggest that assumptions of OLS (the general linear model) were met.
Note 2. r_{sp} denotes semi-partial correlation.
Note 3. VIF denotes the variance inflation factor, which is used to gauge multicollinearity among predictors in the model.

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

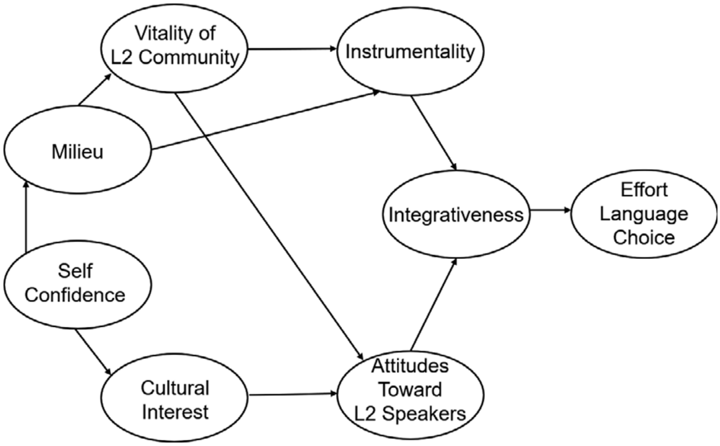

Figure 3. Dörnyei's (2009) proposed L2MSS structure for integrativeness.

(2009) argued that confidence in one's ability leads to the motivational power of the future selves by making them plausible" (ibid., p. 3). The reported semi-partial correlation, r = .28, furthermore, can only be reproduced with this structure (see Table 2). A review of Dörnyei (2009) seems to reveal no such overt claim, however (see Figure 3). Self-confidence is presented as having proximal relationships with "milieu" and "cultural interest" in a multidimensional model with "integrativeness," and then "effort language choice" is the summative outcome(s). Dörnyei does note that integrativeness is a facet of the Ideal L2 Self, but he also was clear that the direct antecedents of the construct did not include Linguistic Self-Confidence (see Figure 2). Papi and Teimouri (2024) also argued that "the two constructs are thus conceptually related, and Linguistic Self-Confidence can be argued to have a close relationship with the Ideal L2 Self (Henry & Liu, 2023)" (p. 3). An automated search of Henry and Liu (2023) for "confidence" and "self-confidence," however, revealed no hits. In a general sense, moreover, the focus of Henry and Liu appears to be on challenging L2MSS as a self-system as opposed to presenting arguments for which constructs in L2MSS theory antecede or predict each other. Figure 1.2 in Henry and Liu does illustrate that possibilities (including can self-possibilities) antecede guides, but this theoretical position is not L2MSS-centric and can self-possibilities, while similar to Linguistic Self-Confidence as a construct, is different in being grounded in the future and is a component of the self-regulatory system, not L2MSS. It therefore becomes apparent that the theory supporting the model producing the reported semi-partial correlations might be lacking according to the conventions of quantitative multivariable inferential research.

Even when giving Papi and Teimouri (2024) the benefit of the doubt in relation to theoretical support for their employed multiple regression model to produce r = .28 (semi-partial correlation), a full consideration of the model challenges their conclusion that, "the analysis showed a modest correlation of between L2 Self-Confidence and Ideal L2 Self, representing nearly 8% of the shared variance and leaving 92% of the variance unexplained. This clearly rejects [Al-Hoorie et al.'s] argument" (p. 4). This is because the semi-partial correlation only tells part of the multiple regression's proverbial story. The underlying F-test for the model, $F(2, 381) = 819.65$, $p < .001$, reveals that 81% (R^2) of the outcome variable's variance has been predicted. Using the logic driving

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

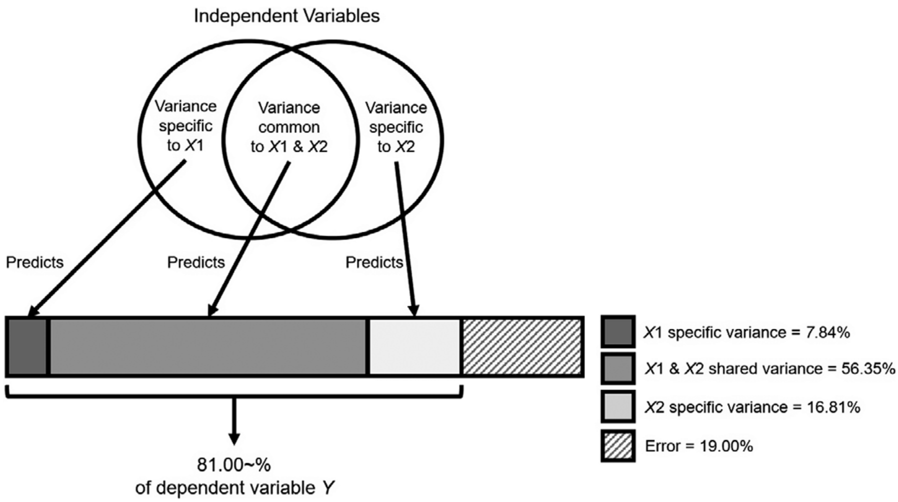

Figure 4. Visualization of the interpretation of Model 1's semi-partial correlations and omnibus multiple regression.
Note: Ideal L2 Self (*Y*), Linguistic Self-Confidence (*X1*), Vividness of Imagery (*X2*).

semi-partial correlations (Abdi, 2007; Aloe, 2014), 7.84% (rounds to 8.00% in Papi and Teimouri [2024]) of this is unique to Linguistic Self-Confidence, and 16.81% of it is unique to Vividness of Imagery, which is visualized in Figure 4. This leaves 56.35% of the variance to be predicted by the common variance between the two predictors in the model. The claim that "92% of the variance is unexplained" is not accurate. On the contrary, only 35.81% of the total variance in the model (i.e., the 16.81% unique to Vividness of Imagery plus the 19.00% unexplained or error variance) cannot be attributed to Linguistic Self-Confidence in some manner using the logic underpinning semi-partial correlations. If 64.19% of the variance is attributable to Linguistic Self-Confidence, either uniquely or in tandem with Vividness of Imagery, then Al-Hoorie et al.'s (2024) concerns about discriminant validity do somewhat hold as it would equate to $r = .80$, which is defensible in the literature as a threshold for collinearity (see discussion in Dormann et al., 2013)[3]. What is more, the observation that most of the explained variance in the model is predicted by the common variance between Linguistic Self-Confidence and Vividness of Imagery further supports Al-Hoorie et al.'s general discriminant validity concerns about L2MSS theory, as two constructs from it cannot account for much of the predicted variance uniquely, just 24.65% as opposed to the 56.35% predicted by the common variance.

3.2 Model 2: Ideal L2 Self (*Y*), Ease of Imagery (*X1*), Vividness of Imagery (*X2*), and Imagery Capacity (*X3*)

Papi and Teimouri (2024) stated that Vividness of Imagery and Ease of Imagery can be considered to be underlying the parental construct of Ideal L2 Self, while Imagery

[3]The threshold of $r \geq .80$ for collinearity appeared to drive Al-Hoorie et al.'s critique and is defensible in the literature and used here. In the interest of fairness, however, one can find support in the literature for more lenient thresholds such as $r \geq .90$ (see Dormann et al., 2013).

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

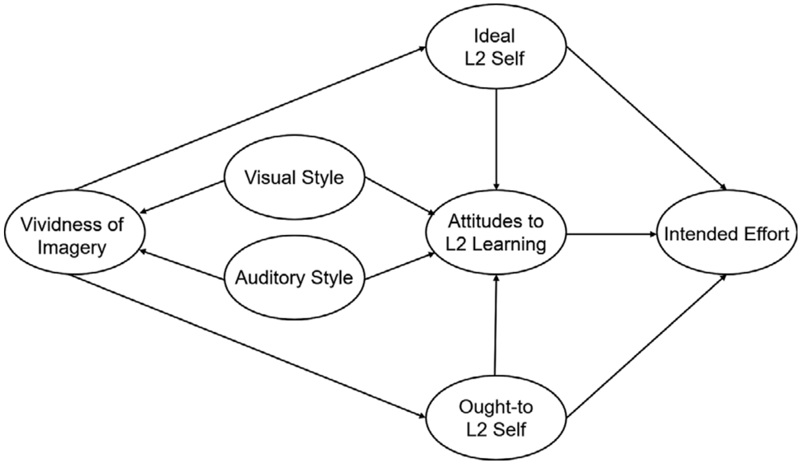

Figure 5. You et al.'s (2016) hypothesized SEM pathways,

Capacity has a predictive causal relationship on Ideal L2 Self. They cited Dörnyei and Chan (2013) who presented regression models where imagery variables predicted measures of the Ideal L2 Self, but Dörnyei and Chan were somewhat explicit that the criterion measure of their study was language achievement. They also claimed that Dörnyei never intended these to be independent constructs. However, You et al. (2016) did not appear to consider Vividness of Imagery to be underlying Ideal L2 Self, but clearly stated that Vividness of Imagery affects both Ideal and Ought-to L2 Future Selves (see Figure 5). Their SEM model clearly displayed a directional predictive relationship of Vividness of Imagery onto the Ideal L2 Self. Furthermore, You et al. (2016. p.116) stated that "Vividness of Imagery strongly affected the two future self-guides—we should note the particularly strong link with the Ideal L2 Self (.81)—which confirms the significant contribution of vividness of mental imagery to future self-guides." We were unable to find any evidence in the work of Dörnyei and colleagues for the relationships between these constructs as described by Papi and Teimouri (2024). This somewhat echoed concerns raised by Hiver and Al-Hoorie (2020) that L2MSS research involving its constructs had contradictory proposed pathways of association (e.g., Taguchi, Magid, & Papi, 2009 [Ideal L2 Self as antecedent] vs. Kormos & Csizér, 2009 [Ideal L2 Self as outcome]). There also appears to be little to no support for using the three constructs related to imagery together in a regression analysis to predict the Ideal L2 Self. You et al. (2016) argued that Vividness of Imagery and Imagery Capacity could be considered to be represented by overall Imagery Capacity.

As with the reconsideration of Model 1, the multiple regression driving Model 2 featured a large omnibus effect where 79% of Ideal L2 Self's variance (R^2) was predicted, $F(3, 380) = 471.26$, $p < .001$. This observation does not fully support Papi and Teimouri's (2024) claims of moderate to weak associations between the variables of interest: "Ideal L2 Self modestly correlated with Vividness of Imagery ($r = .21$), Ease of Imagery ($r = .17$), and Imagery Capacity ($r = .06$) after controlling for correlations among the three variables" (ibid., p. 4). While it is true that only 7.66% of the predicted variance in the model is unique to each predictor (i.e., sum of squared semi-partial correlation values in Table 3, which match what Papi and Teimouri reported), it is

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

Table 3. Abridged reporting of Model 2's (Y: Ideal L2 Self) multiple regression.

Predictor	b*	r_{sp}	VIF
X1. Ease of Imagery	.39	.17	5.29
X2. Vividness of Imagery	.43	.21	4.34
X3. Imagery Capacity	.11	.06	3.16

Note1: Semi-partial correlation values match those reported by Papi and Teimouri (2024).
Note2: Observed Durbin-Watson value (1.92) and residual visualizations suggest that assumptions of OLS (the general linear model) were met.

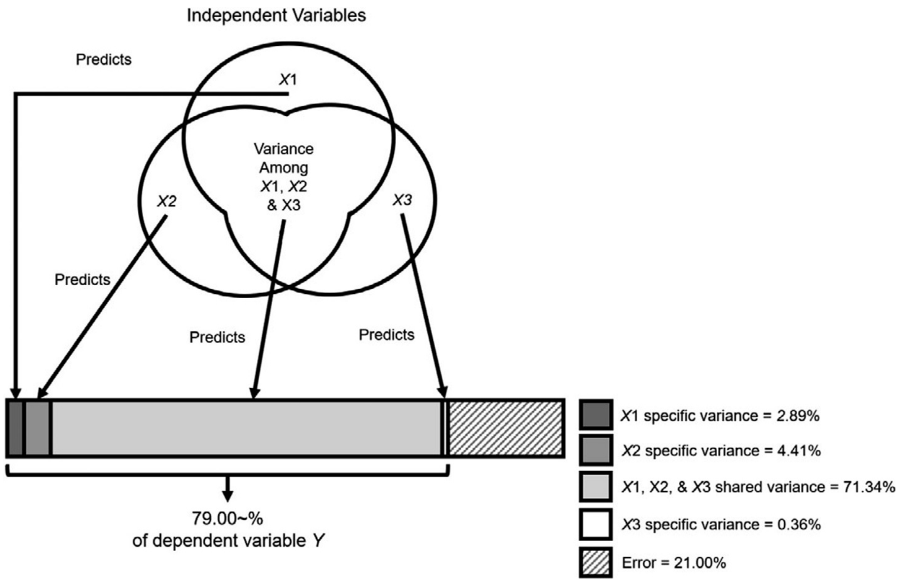

Figure 6. Visualization of the interpretation of Model 2's semi-partial correlations and omnibus multiple regression.
Note: Ideal L2 Self (*Y*), Ease of Imagery (*X1*), Vividness of Imagery (*X2*), and Imagery Capacity (*X3*).

also true that 71.34% of the outcome's variance is predicted by the common (i.e., not unique to one of the predictors) variance among the three predictors, equating to $r = .84$ (see Figure 6). As with the demonstration presented in Figure 2, the amount of total variance that one can ascribe to a predictor using the logic of semi-partial correlations (i.e., unique variance predicting the outcome + common variance among other predictors) also appears to be high enough to suggest co-/multicollinearity and thus discriminant validity concerns are somewhat justifiable. To state it plainly, the common variance among the L2MSS predictor variables is collinear with the outcome variable, Ideal L2 Self. As the predictors themselves are collinear with each other, Al-Hoorie et al.'s (2024) discriminant validity concerns hold. The maximum observed VIF (see Table 3), moreover, only satisfies the most lenient threshold of VIF < 10.00, which some have challenged as being too lenient in avoiding confounding and independence concerns in multiple regression models (e.g., Johnston, Jones, & Manley, 2018). This observation adds to such concerns about measurements of proposed distinct L2MSS constructs not being different when put to the empirical test.

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

Model 3: Ought-to L2 Self (*Y*), Instrumentality-Prevention (*X1*), and Family Influence (*X2*)

Papi and Teimouri (2024) suggested that Instrumentality-Prevention and Family Influence were sibling constructs to the parental construct of Ought-to L2 Self. They then constructed an implicit multiple regression via the reporting of semi-partial correlations where the two sibling constructs were used to predict the Ought-to L2 self. Papi, Bondarenko, Mansouri, Feng, and Jiang (2019), in earlier work, discussed two Ought-to L2 Selves, with one relating to internal beliefs (Ought-to L2 Self-Own: *If I don't work on my English, I will fail in school*) and the other to external beliefs (Ought-to L2 Self-Other: *If I don't learn English, I will disappoint my parents*). This earlier work, additionally, offered no mention of Family Influence and Instrumentality-Prevention as being sibling constructs or potentially being indicators for a governing latent variable. Furthermore, there was no reference to any predictive relationship between them and the overall Ought-to L2 Self as suggested in the proposed regression model in Model 3 of Papi and Teimouri (2024). Henry and Liu (2023) also were cited in support of this model, but their paper contained a strong criticism of the Ought-to L2 Self, arguing that it is unclear how internalized the external pressures are, and also to whom the "others" refer.

Alternatively, Taguchi et al. (2009) featured a SEM where Family Influence and Instrumentality-Prevention were hypothesized to influence the Ought-to L2 Self. Because of Taguchi et al., Model 3 has some theoretical support in the past literature. A deeper examination of Taguchi et al. (2009), however, reveals that the researchers were interested in how these variables influenced the criterion measure, Intended Effort to Study English, as mediated by the Ought-to L2 Self. The model examined the relationship between these variables and Instrumentality Promotion, suggesting bi-directional relationships between these variables. Again, the model does not fully support a regression analysis where these two variables are predictors of the Ought-to L2 Self. Finally, Teimouri (2017) investigated the relationship between the Ought-to L2 Self and Prevention Focus, but bi-directional correlation analyses were used in this study. The only claims of causality (automated search revealed only 1 hit for causal*) in Teimouri (2017) appear to be between Ought-to L2 Self/Own and Ought-to L2 Self/Others:

> Finally, learners' ought-to L2 self/own represents the least extrinsic type of motivation involving external factors that L2 learners have successfully managed to internalize to varying degrees. Its positive correlation with ought-to L2 self/others suggests the extrinsic locus of causality of ought-to L2 self/own, and its stronger correlation with ideal L2 self reflects learners' level of internalization. (p. 700)

As with the first two models, a reconsideration of the multiple regression driving Model 3 does not fully support Papi and Teimouri's (2024) claim that "the results of the analysis showed that Ought-to L2 Self modestly correlated with both Instrumentality-Prevention ($r = .29$) and Family Influence ($r = .37$), supporting distinctions among the measures" (p. 5). The multiple regression accounts for 85% (R^2) of the variance of the outcome variable's Ought-to L2 Self, $F(2, 381) = 1037.36$, $p < .001$. Furthermore, 22.10% of the predicted variance is unique to each predictor using the logic of semi-partial correlations (see Table 4 and Figure 7) with 62.90% of the predicted variance (equating to $r = .79$ - below the collinearity threshold of .80 used by Al-Hoorie et al. [2024])

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

Table 4. Abridged reporting of Model 3's (Y: Ought-to L2 Self) multiple regression.

Predictor	b*	r_{sp}	VIF
X1. Instrumentality-Prevention	.43	.29	2.23
X2. Family Influence	.55	.37	2.23

Note1: Semi-partial correlations match those reported by Papi and Teimouri (2024).
Note2: Observed Durbin-Watson value (1.92) and residual visualizations suggest that assumptions of OLS (the general linear model) were met.

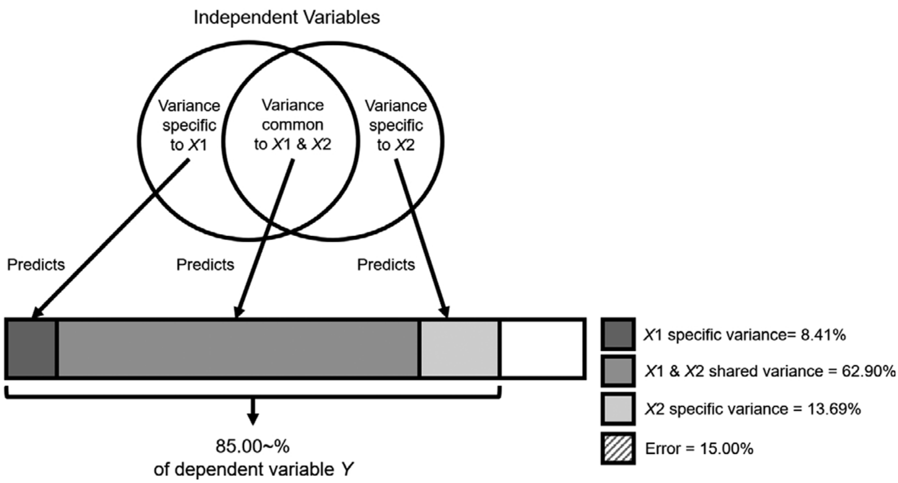

Figure 7. Visualization of the interpretation of Model 2's semi-partial correlations and omnibus multiple regression.
Note: Ought-to L2 Self (*Y*), Instrumentality-Prevention (*X1*), Family Influence (*X2*).

attributable to the common variance between the predictors, as visualized in Figure 7. The collinearity threshold is exceeded, however, when using the logic underlying semi-partial correlations where unique (relative to each predictor) and common variance can be aggregated: Instrumentality-Prevention (71.31% of the predicted variance, $r = .84$) and Family Influence (76.59%, $r = .88$). As reviewed in the preceding text, Abdi (2007) overtly demonstrated how the metric can decompose the omnibus effect of multiple regression in this manner. In sum, Al-Hoorie et al.'s (2024) discriminant validity concerns hold.

Model 4: Intended Effort (*Y*), Attitudes to Learning English (*X1*), Ideal L2 Self (*X2*), and Positive Changes of the Future L2 Self-Image (*X3*)

The final model proposed by Papi and Teimouri (2024) featured Intended Effort as the outcome variable that is predicted by Attitudes to Learning English, Ideal L2 Self, and Positive Changes of the Future L2 Self-Image. The authors cited four sources to support this relationship, but none offer strong support for the implicit multiple regression model.

First, Csizér and Kormos (2009) modeled the motivation of high school and university students, and their model had Learning Experiences (Attitudes to Learning English) and Ought-to L2 Self predicting the Ideal L2 Self, which in turn predicted

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

Motivated Learning Behavior. This does not support the implicit regression model proposed by Papi and Teimouri (2024). Henry and Cliffordson (2017), the second cited source, also modeled data for high school students, but the focus was on discrepancy theory, and they were interested in how the difference between Ideal L2 Self and Current Self predicted Intended Effort. This paper also does not offer support for the proposed regression model.

Taguchi et al. (2009—see Figure 4.2 in their report), the third cited source, constructed a model using L2MSS predictors where Ought-to L2 Self, Attitudes to Learning English, and Ideal L2 Self were the direct antecedents of the criterion (i.e., outcome) measures involving language choice preference and intended effort. As with the preceding cited papers, Taguchi et al. offer little support for Model 4. Finally, You and Dörnyei (2016) conducted a correlation analysis between Intended Effort, Ideal L2 Self, Ought to L2 Self, and Attitudes to L2 Learning. While correlation analyses are underpinned by the same general linear model that drives multiple regression and all other parametric testing, they are considered to be bivariate or bidirectional (Kline, 2023), and when the researchers claim causality using the analysis, they must overtly state so. You and Dörnyei did the opposite when stating "of course, correlations cannot indicate cause–effect relations" (ibid., p. 512), and thus, one can question Papi and Teimouri's (2024) proposed model.

Additionally, Figure 5 illustrates You et al.'s (2016) SEM pathways where Dörnyei was a coauthor. In the strictest sense, this was a multivariable and multivariate model where one hypothesized pathway involved Ideal L2 self is an outcome being predicted by Vividness of Imagery and another pathway featured Intended Effort being predicted by Ideal L2 Self, Attitudes to L2 Learning, and Ought-to L2 Self. In yet another pathway in the model, Attitudes to L2 Learning is predicted by Vividness of Imagery, with its relationship mediated by Ideal L2 Self. These observations lead to further questioning of the strong theoretical support from You and Dörnyei (2016) for Model 4 presented by Papi and Teimouri (2024). In summary, none of the papers cited by Papi and Teimouri (2024) seem to strongly justify a regression model where these variables are used to predict Intended Effort.

The reconsideration of Model 4 as a multiple regression (see Table 5 and Figure 8) challenged the conclusions of Papi and Teimouri (2024) similarly to the previous three models. The multiple regression's omnibus test was significant, $F(3, 380) = 536.52$, $p < .001$, and the effect size was large with 81% (R^2) of the variance explained, where 68.34% of the predicted variance (equating to $r = .83$) of the outcome variable, Intended Effort, was attributable to the common (i.e., not unique to one of the predictors) variance among the three predictors. As the proposed collinearity threshold of .80 was exceeded by the common variance alone, Al-Hoorie et al.'s (2024) discriminant validity concerns were somewhat confirmed. What is more, all VIF values (see Table 5) are above the conservative cutoff of 2.50 (Johnston et al., 2018). While high VIFs are not enough to reject a multiple regression model and strict cutoffs have been challenged in the literature (see

Table 5. Abridged reporting of Model 4's (Y: Intended Effort) multiple regression.

Predictor	b^*	r_{sp}	VIF
X1. Attitudes to Learning English	.45	.29	2.51
X2. Ideal L2 Self	.29	.16	3.28
X3. Positive Changes of the Future L2 Self-Image	.24	.13	3.22

Note1: Semi-partial correlations match those reported by Papi and Teimouri (2024).
Note2: Observed Durbin-Watson value (1.96) and residual visualizations suggest that assumptions of OLS (the general linear model) were met.

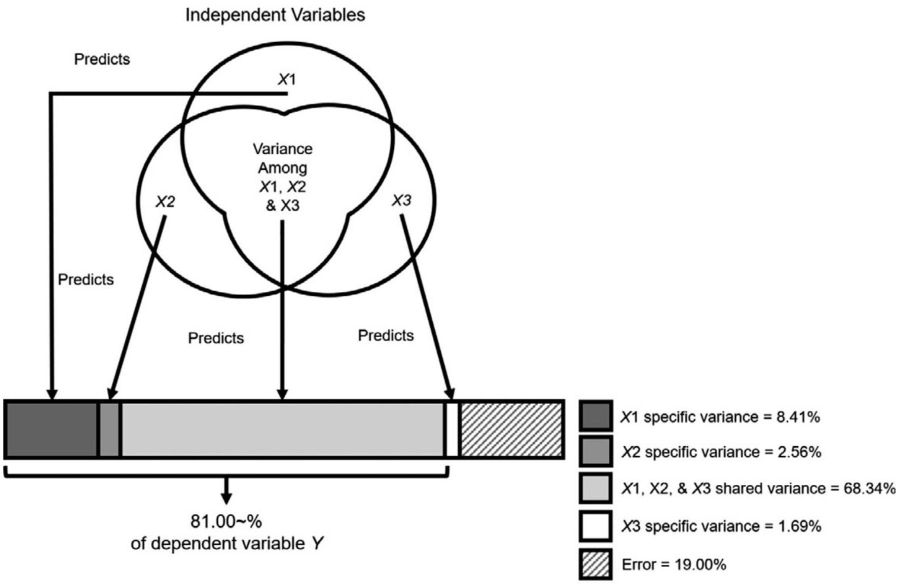

Figure 8. Visualization of the interpretation of Model 4's semi-partial correlations and omnibus multiple regression.
Note: Intended Effort (*Y*), Attitudes to Learning English (*X1*), Ideal L2 Self (*X2*), Positive Changes of the Future L2 Self-Image (*X3*).

e.g., O'brien, 2007), researchers are expected to acknowledge such observations and discuss why they are not invalidating. No such discussion is found in Papi and Teimouri (2024). In summary, the semi-partial correlations' underlying multiple regression for Model 4 might have an unacknowledged independence assumption violation while also confirming Al-Hoorie et al.'s discriminant validity concerns when using the logic underpinning semi-partial correlations.

Concluding remarks

A case for Papi and Teimouri (2024)

Our reconsideration of the semi-partial correlations reported by Papi and Teimouri (2024) contradicts their general position that no collinearity exists among pairs of L2MSS constructs measured by Al-Hoorie et al. (2024). In the interest of criticality and robustness, we consider a two-pronged case for Papi and Teimouri's position despite these observed contradictory findings when considering semi-partial correlations within the multiple regression paradigm. Because of the methods orientation of this response to Papi and Teimouri, we acknowledge but do not engage with their position that definitional validity strengths (Krause, 2012) could supersede discriminant validity weaknesses regarding the L2MSS when presenting this case (for counterpoint expecting the combination of empirical and argument validity evidence; see Campbell & Fiske, 1959; Messick, 1995; Schimmack, 2021). In other words, this presentation of a case for Papi and Teimouri is also exclusively methods-oriented.

The first component of this case for Papi and Teimouri's (2024) position is that they used the measures of Al-Hoorie and colleagues (2024). It is entirely possible that

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

updated and further refined measures might yield different results that support Papi and Teimouri and challenge Al-Hoorie and colleagues. Having stated this, however, Al-Hoorie and colleagues did offer strong arguments for their selection of survey items to operationalize the intended L2MSS constructs. A tangential aspect to this point is that Al-Hoorie and colleagues sampled Korean learners (middle/high school), and it is possible that their findings would not replicate with other populations. In fact, Papi and Teimouri presented lower correlations from past studies conducted in various contexts between L2MSS constructs than those observed by Al-Hoorie and colleagues.

Secondly, Papi and Teimouri's (2024) use of semi-partial correlations appears to have had an unintended consequence relative to their desire to demonstrate low associations among the measurements of interest. Because semi-partial correlations allow for the parsing out of unique and common (more accurately, not unique) variance relative to the predictors and their associations with the outcome variable, the common variance will "dominate" the multiple regression model in terms of explained variance when the predictors are highly correlated. This was the case with their four proposed models. There are other approaches to decomposing R^2 and understanding predictor-level effects (Aloe, 2014; 2015) in multiple regression, such as Pratt's (1987) product measure or relative weight analysis (Nimon & Oswald, 2013), which ascribe the entire amount of variance predicted in the multiple regression model to each predictor. These might have yielded more favorable results for Papi and Teimouri's position. Another possibility could have been communality analysis computations that leverage semi-partial correlations and other multiple regression outputs (Nimon & Oswald, 2013) to further decompose the common variance among 3 or more predictors (e.g., variance common to predictors 1 and 2, but not 3) with which the logic underpinning using semi-partial correlations alone is not directly concerned (see juxtaposition of "commonality analysis" and "squared semi-partial correlation" in Nathans et al., 2012)[4]. As Models 2 and 4 had three predictors, a communality approach could have yielded results more favorable to Papi and Teimouri's position. It bears repeating, however, that addressing discriminant validity concerns and modeling multiple regression with clear outcome and predictor variables are fundamentally different tasks with different approaches proposed for each. Papi and Teimouri (2024) furthermore made no such mention of communality analysis.

Summary

Science makes progress when people question the status quo and engage in lively debate. We welcome the discussion that is currently underway regarding the L2MSS and believe that it will benefit researchers and practitioners who are interested in motivation. Ultimately, the goal of everyone involved in research in the field of motivation is to find ways to positively influence the motivation of language learners, and this kind of discussion and debate is an important part of that process. This paper aimed to describe how semi-partial correlations are historically used as a multiple regression metric, and therefore, caution should be used when interpreting the results

[4]Because Models 1 and 3 (and Abdi's [2007] illustration) only had two predictors, there was only one common variance effect under the logic underpinning commonality analysis (see Nathans et al., 2012). As such, the decomposition of these multiple regression models using semi-partial correlations alone would essentially be viewed as complete commonality analyses.

https://doi.org/10.1017/S0272263125100946 Published online by Cambridge University Press

presented in Papi and Teimouri's (2024) critique of Al-Hoorie et al. (2024). Indeed, the analysis presented in this paper tends to offer support for the original concerns raised by Al-Hoorie et al (2024) of co-/multicollinearity among many of the constructs argued to make up the L2MSS model. We also believe that it is very important to carefully consider theory before conducting any statistical analyses. As we have stated, all research has areas of weakness, and we acknowledge the same for the analyses presented here. We do sincerely hope that this paper makes a positive contribution to the lively discussion, and benefits future researchers, and ultimately the millions of language learners working hard to master another language.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0272263125100946.

Acknowledgements. The authors wish to thank the handling editor, Dr. Luke Plonsky, and the two anonymous reviewers for their constructive feedback during the publication process. Additionally, the authors acknowledge Waseda University's generous support to make this article open access.

Data availability statement. All data used in the report, the aggregation of Al-Hoorie et al.'s (2024, https://osf.io/7c8qs/) data and Abdi's (2007) demonstration data, have been shared via OSF, https://osf.io/vjf6t/. These data allow for the recreation of the visuals, tables, and statistics found in our report.

Competing interests. All authors declare that they have no competing interests.

References

Abdi, H. (2007). Part (semi partial) and partial regression coefficients. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (Vol. *3*, pp. 1–9). SAGE.

Al-Hoorie, A. H., Hiver, P., & In'nami, Y. (2024). The validation crisis in the L2 motivational self system tradition. *Studies in Second Language Acquisition*, Advanced Online Publication. https://doi.org/10.1017/S0272263123000487

Aloe, A. M. (2014). An empirical investigation of partial effect sizes in meta-analysis of correlational data. *Journal of General Psychology*, *141*(1), 47–64. https://doi.org/10.1080/00221309.2013.853021

Aloe, A. M. (2015). Inaccuracy of regression results in replacing bivariate correlations. *Research Synthesis Methods*, *6*(1), 21–27. https://doi.org/10.1002/jrsm.1126

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates Publishers.

Csizér, K., & Kormos, J. (2009). Learning experiences, selves and motivated learning behavior: A comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei & E. Ushioda, (Eds.), *Motivation, language identity and the L2 self* (pp. 98–119). Multilingual Matters.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G. … Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.

Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self images, sensory styles, and imagery capacity across two target languages. *Language Learning*, *63*(3), 437–462. https://doi.org/10.1111/lang.12005

Dunlap, J. W., & Cureton, E. E. (1930). On the analysis of causation. *Journal of Educational Psychology*, *21*(9), 657–680. https://doi.org/10.1037/h0071119

Engelhart, M. D. (1936). The technique of path coefficients. *Psychometrika*, *1*, 287–293. https://doi.org/10.1007/BF02287881

Eguchi, M., & Kyle, K. (2020). Continuing to explore the multidimensional nature of lexical sophistication: The case of oral proficiency interviews. *Modern Language Journal*, *104*(2), 381–400. https://doi.org/10.1111/modl.12637

Henry, A., & Cliffordson, C. (2017). The impact of out-of-school factors on motivation to learn English: Self-discrepancies, beliefs, and experiences of self-authenticity. *Applied Linguistics*, *38*(5), 713–736. https://doi.org/10.1093/applin/amv060

Henry, A., & Liu, M. (2023). Can L2 motivation be modelled as a self-system? A critical assessment. *System*, *119*, 103158. https://doi.org/10.1016/j.system.2023.103158

Henry, A., & Liu, M. (2024). Jingle–jangle fallacies in L2 motivational self system research: A response to Al-Hoorie et al. (2024). *Applied Linguistics*, *45*(4), 738–746. https://doi.org/10.1093/applin/amae041

Henson, R., & Roberts, J. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*, 393–416. doi: 10.1177/0013164405282485

Hiver, P., & Al-Hoorie, A. H. (2020). Reexamining the role of vision in second language motivation: A preregistered conceptual replication of You, Dörnyei, and Csizér (2016). *Language Learning*, *70*(1), 48–102. https://doi.org/10.1111/lang.12371

Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, *28*(1), 201–230. https://doi.org/10.1177/13621688211001289

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality & Quantity*, *52*(4), 1957–1976. https://doi.org/10.1007/s11135-017-0584-6

Krause, M. S. (2012). Measurement validity is fundamentally a matter of definition, not correlation. *Review of General Psychology*, *16*(4), 391–400. https://doi.org/10.1037/a0027701

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.

Kormos, J., & Csizèr, K. (2009). Learning experiences, selves and motivated learning behaviour: A comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 98–119). Multilingual Matters.

Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: indices, tools, findings, and application. *TESOL Quarterly*, *49*(4), 757–786. https://doi.org/10.1002/tesq.194

Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, *34*, 12–24. https://doi.org/10.1016/j.jslw.2016.10.003

Larsson, T., Plonsky, L., Sterling, S., Kytö, M., Yaw, K., & Wood, M. (2024). On the frequency, prevalence, and perceived severity of questionable research practices. *Research Methods in Applied Linguistics*, *2*(3), 100064. https://doi.org/10.1016/j.rmal.2023.100064

Lawson, K. M., & Robins, R. W. (2021). Sibling constructs: what are they, why do they matter, and how should you handle them? *Personality and Social Psychology Review*, *25*(4), 344–366. https://doi.org/10.1177/10888683211047101

Leeming, P., Vitta, J. P., Hiver, P., Hicks, D., McLean, S., & Nicklin, C. (2024). Willingness to communicate, speaking self-efficacy, and perceived communicative competence as predictors of second language spoken task production. *Language Learning*, *74*(4), 777–1025. https://doi.org/10.1111/lang.12640

McNemar, Q. (1962). *Psychological statistics* (3rd *ed.*). Wiley.

Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement Issues and Practice*, *14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research & Evaluation*, *17*(9), 1–19. https://doi.org/10.7275/5fex-b874

Nimon, K. F., & Oswald, F. L. (2013). Understanding the results of multiple linear regression: Beyond standardized regression coefficients. *Organizational Research Methods*, *16*, 650–674. https://doi.org/10.1177/1094428113493929

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Oga-Baldwin, W. L. Q. (2024). Validation crisitunity. *Studies in Second Language Acquisition*, Advanced Online Publication. https://doi.org/10.1017/s0272263124000597

Papi, M., Bondarenko, A. V., Mansouri, S., Feng, L., & Jiang, C. (2019). Rethinking L2 motivation research: THE 2 × 2 model OF L2 self-guides. *Studies in Second Language Acquisition*, *41*(2), 337–361. doi:10.1017/S0272263118000153

Papi, M., & Teimouri, Y. (2024). Manufactured crisis. *Studies in Second Language Acquisition*, Advanced Online Publication. https://doi.org/10.1017/s0272263124000494

Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. *Studies in Second Language Acquisition*, *39*(3), 579–592. https://doi.org/10.1017/s0272263116000231

Pratt, J. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukkila & S. Puntanen (Eds.), *Proceedings of the Second International Tampere Conference in Statistics* (pp. 245–260). University of Tampere.

Schimmack, U. (2021). The validation crisis in psychology. *Meta-Psychology*, *5*, 1–9. https://doi.org/10.15626/mp.2019.1645

Taguchi, T., Magid, M., & Papi, M. (2009). The L2 motivational self system among Japanese, Chinese and Iranian learners of English: A comparative study. In Z. Dörnyei, Z. & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 66–97). Multilingual Matters.

Teimouri, Y. (2017). L2 selves, emotions, and motivated behaviors. *Studies in Second Language Acquisition*, *39*(4), 681–709. https://doi.org/10.1017/s0272263116000243

Velicer, W. F. (1978). Suppressor Variables and the Semipartial Correlation Coefficient. *Educational and Psychological Measurement*, *38*(4), 953–958. https://doi.org/10.1177/001316447803800415

Vitta, J. P., Nicklin, C., & Albright, S. W. (2023). Academic word difficulty and multidimensional lexical sophistication: An English-for-academic-purposes-focused conceptual replication of Hashimoto and Egbert (2019). *Modern Language Journal*, *107*(1), 373–397. https://doi.org/10.1111/modl.12835

Worthington, R., & Whittaker, T. (2006). Scale development research: A content analysis and recommendations for best practices. *Counseling Psychologist*, *34*, 806–838. https://doi.org/10.1177/0011000006288127

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *20*, 557–585.

You, C., & Dörnyei, Z. (2016). Language learning motivation in China: Results of a large-scale stratified survey. *Applied Linguistics*, *37*(4), 495–519. https://doi.org/10.1093/applin/amu046

You, C., Dörnyei, Z., & Csizèr, K. (2016). Motivation, vision, and gender: A survey of learners of English in China. *Language Learning*, *66*, 94–123. https://doi.org/10.1111/lang.12140

Cite this article: Vitta, J. P., Leeming, P., & Nicklin, C. (2025). Interpreting the semi-partial correlation as a multiple regression-bound (not a bivariate) metric: A methods-oriented response to Papi and Teimouri's (2024) response to Al-Hoorie et al. (2024). *Studies in Second Language Acquisition,* 47: 1148–1169. https://doi.org/10.1017/S0272263125100946