5-11-2020

Technical Report for "When People Estimate their Personal Intelligence Who Is Overconfident? Who is Accurate?"

John D. Mayer
University of New Hampshire, jack.mayer@unh.edu

A. T. Panter
University of North Carolina

David R. Caruso
Yale University

Technical Report for "When People Estimate their Personal Intelligence Who Is Overconfident? Who is Accurate?"

John D. Mayer

University of New Hampshire

A. T. Panter

University of North Carolina

David R. Caruso

Yale University

Contents

This page intentionally left blank.

Introduction to the Technical Supplement

The Technical Supplement includes additional information about the article "Who Believes they are High in Personal Intelligence." The Supplement is organized such that material follows the organization of the article, with the exception that group-wise analyses—i.e., analyses based on median splits of the archival samples on the Test of Personal Intelligence and Self-Estimated Personal Intelligence, are in their own Appendix owing to the considerable length of that material.

Disclaimer: Please note also that in order to provide continuity and transitions between the main document and this technical supplement, and because the documents were developed in tandem by the same authors, some (usually brief) sections from the article may be included verbatim or almost so in this Supplement, and may not appear in quotes.

Supplement to "Archival Study Methods"

The impetus for the present project was the question of how people who were accurate versus inaccurate regarding estimates of their own personal intelligence might differ from one another. To address this question systematically, we identified samples we had collected in our lab for which participants had completed versions of both the Test of Personal Intelligence (TOPI) and of the measure of Self-Estimated Personal Intelligence (SEPI), and some other criterion measure(s) of interest. (We possessed other data sets that contained just the TOPI and SEPI with no further psychological measures, as they had been collected to refine the measurement instruments themselves).

We identified seven potentially relevant data sets collected between 2012 and the present (excluding collaborations with others outside the lab and excluding projects led by graduate students), of which three such studies included both measures, indicated in the Overview Table 1.

For those samples included, we used the data set as it was employed in any published report: That is, we followed the original method of screening and employed the unaltered data set.

Supplemental Table 1

Inclusion Criteria and Selection of Studies

Manuscript or Publication	Approx. Yrs of Data Collection	TOPI Version	SEPI version	Big Five?	Other Comments
1. Does PI Exist? (Mayer et al., 2012)	2010-2012	2	None	Yes	Prior to SEPI development
2. Alternate Measures data (Mayer et al., n.d.)	2011-2013	MINI-12	SEPI-120	Yes	Unpublished
3. A closer look at PI using the TOPI (Mayer et al., 2017)	2013-2017	4, 4R	SEPI-16	No	Uses archival data on TOPI only
4. Cadets at West Point (Mayer & Skimmyhorn, 2017)	2013-2014	2R	SEPI-16 (not included in analyses)	Yes	The SEPI-16 data went unexamined/unanalyzed, the data are held by the OEMA, and we have lost access to it short of a great deal of effort
5. Employees high in personal intelligence…Study 1 (Mayer et al., 2018)	2014-2015	MINI-12	SEPI-74	No	…but other criteria scales of interest
6. Employees high in personal intelligence…Study 2 (as above)	2015-2017	4R	SEPI-16	Yes	
7. Advancing the measurement of personal intelligence (Mayer et al., 2019)	2015-2017	5 and 5R	None	No	psychometric development only

text between tables

section break next page

Supplement to "Data Analytic Strategies"

1. A first approach involves a median split approach, in which we create four groups: those high or low in personal intelligence crossed with those who estimated their abilities as high or low (i.e., low-low, low-high, high-low, and high-high. This method is often helpful in applied settings when setting cutpoints for selection, for example. The same method is often criticized, however, because of its relatively low power compared to creating continuous variables, and the possibility that by moving cutpoints slightly, the significance of results may be altered (DeCoster et al., 2011; MacCallum et al., 2002).

2. A second method involves the creation of differences scores between self-estimated ability and actual ability. This approach has the strength of providing continuous data, but is often criticized owing to the frequent unreliability of difference scores (MacCallum et al., 2002).

3. A third approach involves the use of residual scores in which self-estimates are predicted from ability scores, and the residuals are calculated as an index of a person's self-enhancement (or lack thereof). They represent the bias that remains after the "reality" component has been partialed out (John & Robins, 1994). This drawback of this approach is that can leave a high correlation between the residuals and the original self-estimates (Krueger et al., 2017).

4. Yet a further approach is polynomial regression with response surface analysis (RSA) (Edwards, 2009).

From our perspective, all these approaches have their merits and limits, and the choice of the best technique is, to a degree, an issue of the specific question being asked and the nature of the data being analyzed. In the present case, we centered our analyses on the use of continuous variables using differences between the z-scores (DIZs, Laird & Weems, 2011) of estimated and actual ability. That said, polynomial-with-RSA analyses may be found in Appendix A and median split groups in Appendix B.

Supplement to "Calculation of Accuracy and Overconfidence Scores"

Difference scores remain a tool of choice, especially difference scores from z-scores or DIZs, which are placed on the same metric (Laird & Weems, 2011, p. 389). In Trafimow's (2015) recent defense and reexamination of difference scores, revisiting its origins in classical test theory, he asks whether such scores are truly unreliable and answers, "it depends"—if the two tests are reliable and their intercorrelation is modest, as in the present instances, then some reliability is apt to be present (Trafimow, 2015, p. 4). There appeared to be at least a promise that would be the case regarding the present data, given the relatively low correlations among the TOPI and the SEPI.

The reliability of a difference score is calculated in one form (assuming equal variances, which follows from using DIZs (i.e., equalizing standard deviations) is:

$$p_{diff} = \frac{p_{xx'} + p_{yy'} - 2p_{xy}}{2 - 2p_{xy}}$$

(Linn & Slinde, 1977, p. 123, Formula 2, simplified for the case where S = 1).

In our present work, working through the reliabilities for each study, we get the following:

text between tables

Supplemental Table 2

Reliability of difference scores

	Sample reliabilities of original measures			Correlation between measures		Numerator $p_{xx'} + p_{yy'} - 2p_{xy}$	Denominator $2 - 2p_{xy}$	Rel. of diff.
	TOPI	SEPI	Sum	As is	Times two			
Study 1	.64	.83	1.47	.06	.12	1.35	1.88	.72
Study 2	.71	.89	1.60	.11	.22	1.38	1.78	.78
Study 3	.94	.95	1.89	.28	.56	1.33	1.44	.92

text between tables

The related formula for the reliability of the residuals is:

$$p_{diff} = \frac{p_{yy'} - p_{xy'}^2(2 - p_{xx'})}{1 - p_{xy'}^2}$$

(Linn & Slinde, 1977, p. 125, Formula 4).

text between tables

Supplemental Table 3

Reliabilities of Accuracy Residuals for ***TOPI*** as Y Variable Across Studies

	Sample reliabilities		Correlation between measures		$(2 - p_{xx'})$	$p_{xy'}^2(2 - p_{xx'})$	Numerator $p_{yy'} - p_{xy'}^2(2 - p_{xx'})$	Denominator $1 - p_{xy'}^2$	Rel. of res.
	TOPI (as Y)	SEPI (as X)	$r_{SEPI \cdot TOPI}$	Squared	(compute)				
Study 1	.64	.83	.06	.0036	1.17	.0042	.636	.996	.64
Study 2	.71	.89	.11	.0121	1.11	.0134	.697	.988	.71
Study 3	.94	.95	.28	.0784	1.05	.0823	.857	.922	.93

text between tables

Supplemental Table 4

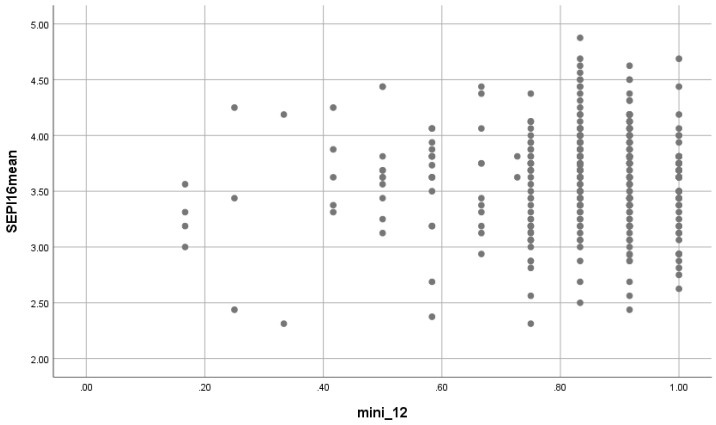
Reliabilities of Confidence Residual for ***SEPI*** as Y Variable Across Studies

	Sample reliabilities		Correlation between measures		$(2 - p_{xx'})$	$p_{xy'}^2(2 - p_{xx'})$	Numerator $p_{yy'} - p_{xy'}^2(2 - p_{xx'})$	Denominator $1 - p_{xy'}^2$	Rel. of res.
	TOPI (as X)	SEPI (as Y)	$r_{SEPI \cdot TOPI}$	Squared	(compute)				
Study 1	.64	.83	.06	.0036	1.36	.0048	.825	.996	.86
Study 2	.71	.89	.11	.0121	1.29	.0156	.874	.988	.89
Study 3	.94	.95	.28	.0784	1.06	.0831	.867	.922	.94

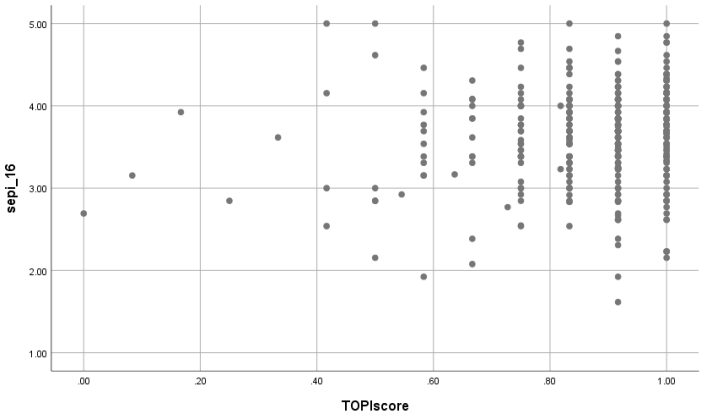
- Bennett, J., & Briggs, W. (2005). *Using and understanding mathematics: A quantitative reasoning approach* (3rd ed.). Boston: Pearson.
- Törnqvist, L.,Vartia, P., Vartia, Y. (1985). How should relative changes be measured?, *The American Statistician*, 39, 43–46.

Supplement to "The Basic Scores": SEPI and TOPI Scatterplots Across the Three Studies

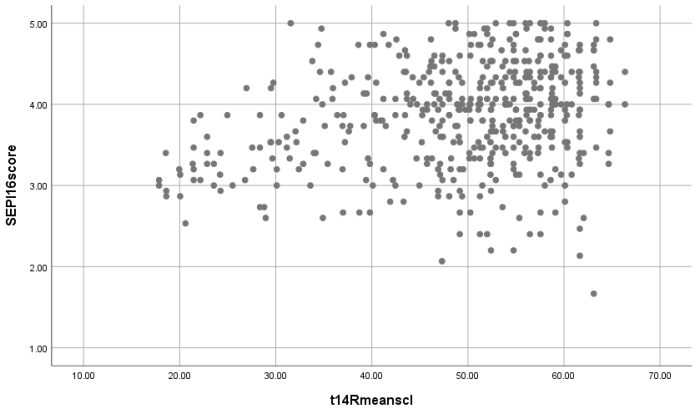
Study 1



Study 2



Study 3



Supplement to "General Checks" Concerning Differences Between Women and Men

In Studies 1 and 3, women exhibited both higher personal intelligence and higher estimates of their personal intelligence than men; there was no difference in Study 2. The differences were small, i.e., on the order of .25 standard deviations, but present. This is the reverse of the sometimes-reported trend that men estimate their intelligence more highly than women. Perhaps women "own" the people-centered intelligences on average, and men acknowledge that ownership on average (though there are many individual exceptions). Turning to the scores for accuracy and overconfidence, there were no consistent differences for any of the scores across studies; only two of the twelve differences tested were significant at all.

buffer text for table

Supplemental Table 5

Means and Standard Deviations of Basic and Accuracy-Adjusted Scores by Women and Men Across Studies

					Accuracy		Confidence	
		N	SEPI	TOPI	DIZ	Residual	DIZ	Residual
Study 1	Men	97	3.52	0.79	50.60	49.05	50.18	47.85
	Women	247	3.66	0.85	49.91	50.49	49.74	50.74
	t_{diff}		-2.69**	-3.18**	.573	-1.22	.367	-2.44*
Study 2	Men	176	3.65	0.87	50.00	49.15	50.34	49.99
	Women	216	3.66	0.88	49.98	50.67	49.69	49.98
	t_{diff}		-.092	-.932	.023	-1.503	.638	.009
Study 3	Men	230	3.79	46.92	50.24	49.03	50.69	49.28
	Women	250	3.95	51.29	49.79	50.89	49.43	50.70
	t_{diff}		-2.75**	-.466***	.494	-.203*	1.38	-1.56
Note: t values for independent *t*-tests assuming equal variances								

buffer text around table

Supplement to "Response Surface Analyses of Accuracy Scores"

Prolegomenon.

A further possibility for assessing congruence effects is polynomial regression coupled with Response Surface Analysis (RSA), which has become an increasingly popular approach to this research question. Proponents of the technique argue that (a) polynomial regression tends to fit better than linear regression, and that the two variables are treated separately in the RSA such that one can see the specific action of each variable

in the surface diagram. Accordingly, we conducted a number of analyses of our data using polynomial regression with response surface analyses (RSA), as well as the analyses reported in the published report.

Here we briefly report the results of those analyses, with interpretations of the figures we present. Those unfamiliar with the technique, we hope, nonetheless will be able to understand the graphs we present and their interpretations. That said, the bases for understanding these techniques require some explanation beyond the scope of this technical supplement. For key resources, the interested reader can refer to a number of good published sources (e.g., Edwards, 2002; Humberg et al., 2019; Shanock et al., 2010).

Limitations of the Technique.

It is worth mentioning that polynomial regression with RSA (we often speak simply of RSA) is recognized as having a number of limits to its application as conventionally employed. These include that (a) as more model parameters are estimated in the context of a single regression than is typical, the statistical power of the approach may be reduced, (b) those who employ the technique "assume all variables have been measured without error", i.e., with perfect reliability (e.g., Schönbrodt et al., 2018, p. 638), (c) quadratic formulae—the conventionally used level of polynomial regression—forces symmetry on data which, in fact, may not be symmetrical (cf. Humberg et al., 2019, on enforced symmetry), and (d) the technique does not generate scores analogous to DIZs or residuals, for example, that can be used for further analyses.

These limits affected our results to greater and lesser (and clearer and less-clear) degrees, as we will describe below. First, however, we present the results of our analyses.

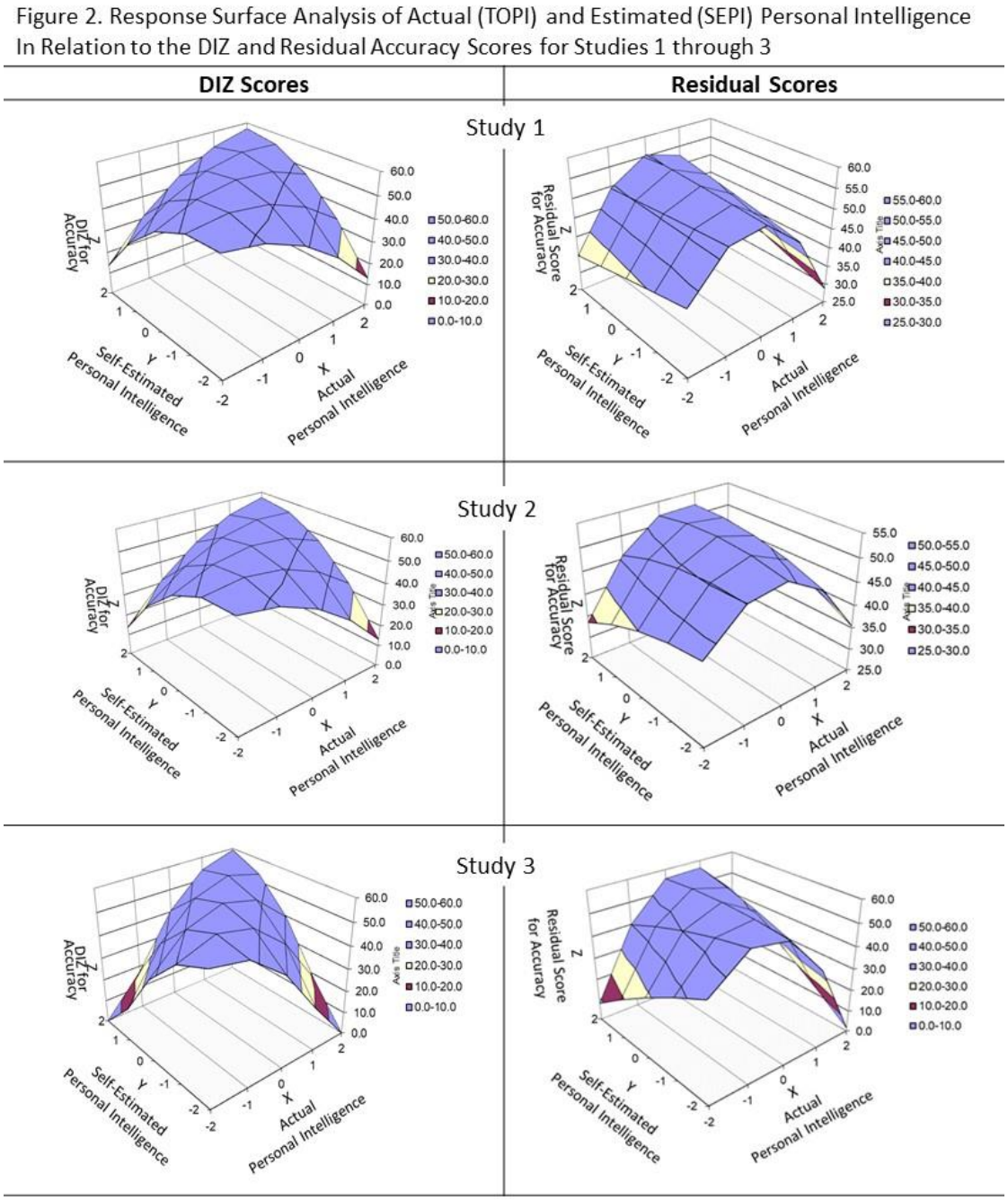
Analyses Conducted

Recall that our report is exploratory, and consequently, we focused on modeling only those findings that *a posteriori* were of theoretical interest or, alternatively, analyses relating the SEPI and TOPI to correlates that had shown some theoretically and empirically meaningful relation across studies. The scope of our applications were to: (a) explore the performance of the DIZ and residual scores we had formed, to (b) relate conscientiousness and vocabulary from the SEPI and TOPI and (c) to relate counterproductive work behavior to the SEPI and TOPI.

Performance of the DIZ and Residual Scores Through the Lens of Response Surface Analysis

Response surface analysis promises to allow researchers to view how two variables affect the level of a third. One of our applications, then, was to see how the SEPI and TOPI affected DIZ scores and residual scores. If our conceptualization and calculation of DIZs were correct, for example, we would expect the RSAs graphs—which are drawn in three dimensions—would resemble the shape of a parafoil (or of a skate with its pectoral flaps down), with the spine of the parachute (or animal) oriented at a 45-degree angle with the highest points along a line where X = Y, that is, the accuracy DIZs would be highest where the SEPI equaled the TOPI (with everything on T-scales). This is exactly what the RSA analyses yielded for all three studies, as indicated in the graphs depicted in what is labeled as Appendix A Figure 1 (and below that label, as Figure 2). In all studies, the a4 parameter was negative, and statistically significant so, allowing one to conclude that there was a statistically significant congruence effect (Humberg et al., 2019).

Appendix A Table 1 Response Surface Analysis of Actual (TOPI) and Estimated (SEPI) Personal Intelligence in Relation to DIZ and Residual Accuracy Scores for Studies 1 through 3.



Note: All polynomials regression *R*s between .91 and .95. The parafoil shape to the left reflects a near-perfect convergence effects for the Difference-in-Zs. The effects to the right reflect convergence effects as well, but exhibit greater influence from the TOPI compared to the SEPI, suggesting some undercorrection. All A4 terms were stastistically significant as were most A3 terms (see Humberg, Nestler, & Back, 2019).

Moreover, the results for the residual scores for accuracy were quite similar to those for the DIZs—except that this time, the spine of the figures were imperfectly oriented along the X = Y axis. The rotation favoring the TOPI indicated that, as many psychometricians have pointed out, residual scores often are undercorrected

relative to DIZs: that is, they are more saturated with the variance from one of their composites than the other (for these accuracy residuals, the TOPI).

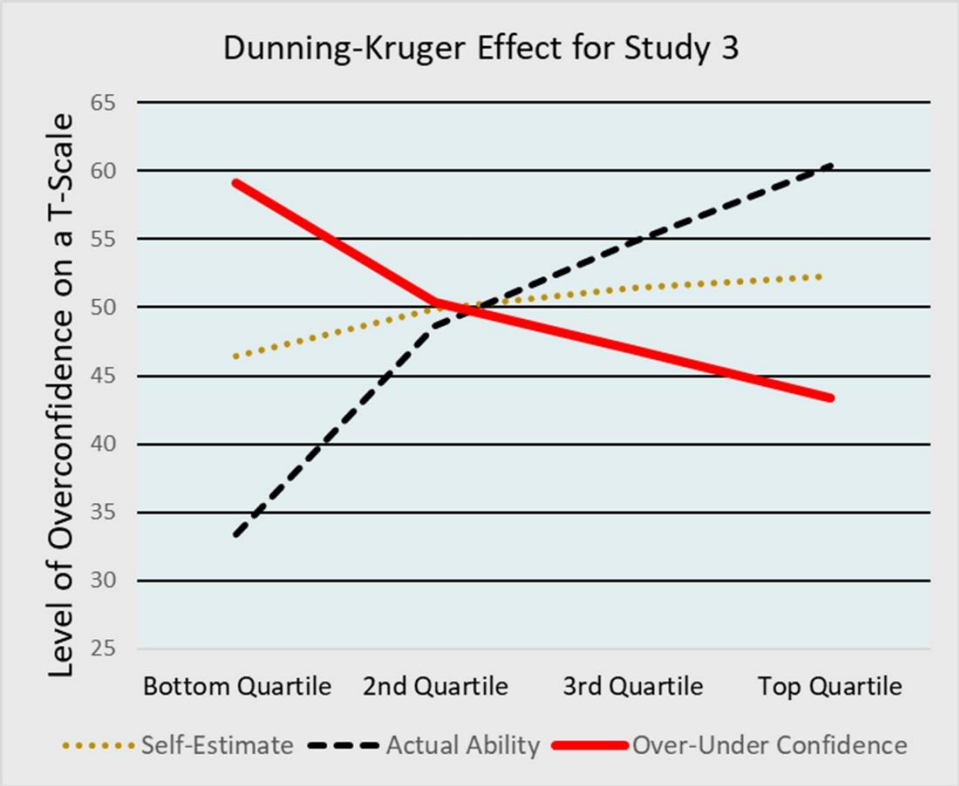
Supplement to "Dunning-Kruger Effects"

We also conducted tests for a Dunning-Kruger effect: That people who are least able in ability will express the highest degree of overconfidence relative to the other groups. The analysis customarily involves dividing the sample into four quartiles of ability and then conducting a test for over-underconfidence (i.e., the self-estimate minus the ability level, both converted to the same scale) for differences across the four groups.

In Studies 1 and 2, the TOPI-MINI was used, and because of its limited score range, our interquartile cutpoints ended up merging two quartiles (2 and 3) in Study 1 and again in Study 2 (3 and 4). All four quartiles could be distinguished in Study 3 which used a full-length TOPI form. Note that the lowest quartile could be separated out in all three and we obtained Dunning-Kruger effects across all three studies. Tukey's Honest Significant Difference (HSD) tests indicated that Group 1 (the lowest ability group) was most overconfident in Study 1 (M_{Q1}= 57.87 versus $M_{Q2,3}$ = 48.85 and M_{Q4} = 41.09) all groups different < .05. In Study 2, Group 1 again was most overconfident (M_{Q1}= 60.79 versus M_{Q2} = 49.34 and $M_{Q3,4}$ = 44.45), all groups different $p < .05$. And the same held for Study 3 for the longer TOPI form (M_{Q1}= 59.16 versus M_{Q2} = 50.39, M_{Q3} = 46.94, *and* M_{Q4} = 43.36), all groups different $p < 05$ for Tukey's HSD. The effect can be depicted as in Appendix D Figure 1.

Appendix D Figure 1

Overconfidence is Highest Among the Least Skilled: Results from Study 3



x

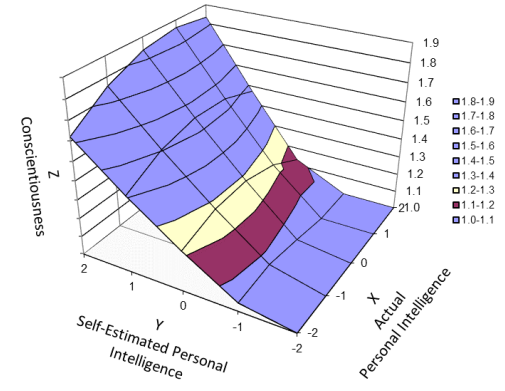
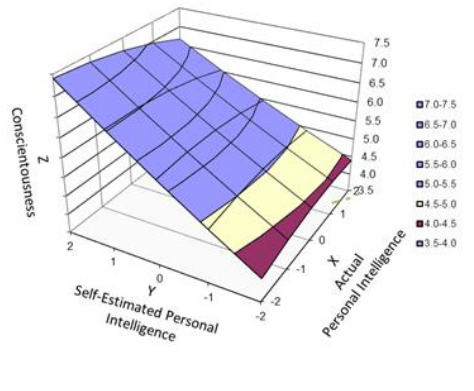
Supplement to "Applying RSA to Vocabulary and Conscientiousness"

Response Surface Analysis in Studies 1 and 3 for Conscientiousness

The RSA results of the TOPI and SEPI, as associated with conscientiousness in Study 1 fit well but was unremarkable and limited. The surface plot for conscientiousness (Figure X.1) exhibits a simple plane—a flat surface (approximately) here and, later in Study 3 (to the right), indicating simply that conscientiousness rose with the SEPI scores: The only significant term in the polynomial regression (aside from the overall *R*) in Study 1 was the SEPI main effect ($B = .26$ $p < .001$), and the same for Study 3 ($B = .695$ $p < .001$). That said, there is a hint of negative "spinal" curve—but little sign of it statistically.

Recall that conscientiousness correlated with greater accuracy in personal intelligence estimates according to DIZ and residual scores, and also that conscientiousness was closely related to self-estimated personal intelligence in Studies 1 and 3. Yet in the polynomial regression, neither the squared (quadratic) or interaction terms were significant, and so the response surface manifested more-or-less as an angled plane (e.g., Humberg, Nestler & Black, 2019, p. 10, right col.): When self-estimates rose, so did conscientiousness somewhat apart from the influence of actual personal intelligence.

Appendix A. Figure 2
Self-Judged and Objective Personal Intelligence, with Conscientiousness From Studies 1 and 3

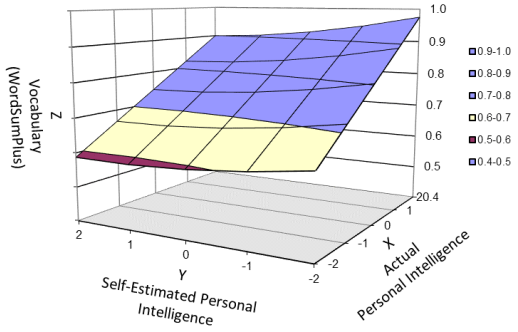
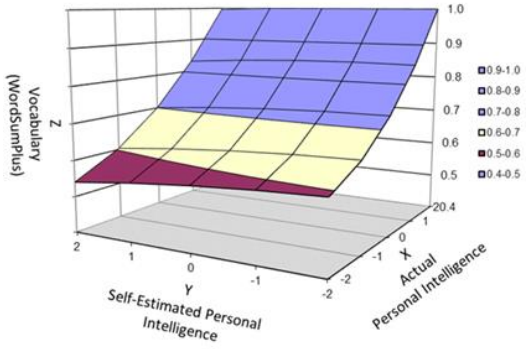
Study 1 Conscientiousness	Study 3. Conscientiousness
	

The SEPI-TOPI RSA in Study 2 for Vocabulary

The RSA results of the TOPI and SEPI, as associated with vocabulary in Study 2 was again unremarkable. The surface plot for vocabulary (Figure Xb) again exhibited a simple plane, indicating simply

that vocabulary rose with TOPI scores (B = .08 p < .001), with a much slighter negative contribution from the SEPI (B = -.015, p < .1).

Appendix A. Figure 3
Self-Judged and Objective Personal Intelligence, with Vocabulary From Studies 2 and 3

Study 2 Vocabulary	RSA Study 3 Vocabulary
	

The results from Study 3 was much the same as in Study 2: The surface plot for vocabulary (Appendix A, Figure 2) again exhibited a simple plane, indicating simply that vocabulary rose with TOPI scores (B = .11 p < .001), with a much slighter contribution from the term representing TOPI-squared (B = -.017, p < .05). when actual PI rose, so did vocabulary. There was, however, also an accelerating curve to the plane for the relation between personal and verbal intelligences in Study 3, indicated by a significant beta for the squared TOPI score (see Appendix A, Table 1). Perhaps people with more personal intelligence better recognize the value of understanding the language, or those with larger vocabularies better "speak the language" of personal intelligence.

Supplement to "Tests for Interaction of Counterproductive Work Behavior in Study 2 and Study 3"

Study 2

Beginning statistical setup needed for the analyses

Note: Actual information and numbers to input are highlighted blue; the calculated values are highlighted yellow

Description of or Title for Data: Additional analysis of Study 2 from Mayer, Lortie, Panter & Caruso, 2018

Step 1. Identify the necessary variables **Step 2.** Enter the specific values **Step 3.** Review the calculations and read the figure from the "Plot" pag

Regression Relevant Information				General Statistical Info			
Variable	Description	Name	Coefficient	B	Mean	Std. Dev	Coeff. Variances
Y	Dependent	CWB			1.2435	0.2986	0.08916196
Intercept	Crosses Y a	--	b_0	1.293	--	--	
X	predictor 1	TOPI	b_1	-0.042	0	1	1
Z	predictor 2	SEPI	b_2	-0.067	0	1	1
XZ	Interaction	iTOPISEPI	b_3	0.029	1.07	1.21	1.4641

Values of X desired for the interaction lines to begin and end: -1.5 1.5 *Note:* -1.5 and 1.5 to distinguish from the Z values

We know the mean and standard deviation of Z, so to compute the value 1 SD below, at the mean and 1 SD above the mean is fairly straightforward. The values are:

	Low	Medium	High
Values of Z (Predictor 2)	-1	0	1

Note: These values are set using the M and S of Z as +/- 1 SD from M

Relevant Starting Equations

The generalized formula for regression, with a bivariate interaction term is: These are equations

$$Y_{pred} = b_0 + b_1X + b_2Z + b_3(X*Z) + e$$

These, then, are rearranged to focus on one predictor at a time, e.g., X as a predictor:

$$Y_{pred} = (b_1 + b_3Z)X + (b_0 + b_2(Z))$$

The above equation has two principle terms designated as indicated below into the simple slope and simple intercept:

The Simple Slope	The Simple Intercept
$(b_1 + b_3Z)$	$(b_0 + b_2(Z))$

Using the formulae above, and taking the starting values entered initially, these become

To find the regression line, we must set the values of Z and solve for X (we also could set X and solve for Z, but that isn't our objective for the figu

Conventionally, the values of Z are set at a low, medium, and high value, and then we solve for the Xs that will define the line relative to the prec

	The Simple Slope	$(b_1 + b_3Z)$	Note: X is *not* part of the simple slope--that comes later					SIMPLE SLOPE
	Z Variable	(b_1	plus	(b_3	times	Z	)	
Low Z	SEPI	(-0.042	plus	(0.029	times	-1)	equals	-0.096
Medium Z	SEPI	(-0.042	plus	(0.029	times	0)	equals	-0.067
High Z	SEPI	(-0.042	plus	(0.029	times	1)	equals	-0.038

	The Simple Intercept $(b_0 + b_2(Z))$							
	(b_0	plus	(b_2	times	Z	)		SIMPLE INTERCEP
Low Z	SEPI	1.293		-0.042		-1	equals	1.335
Medium Z	SEPI	1.293		-0.042		0	equals	1.293
High Z	SEPI	1.293		-0.042		1	equals	1.251

So...using the formulae above, the simple slope and simple intercept for the three lin Values of X (taken from row 11: "Values of X at which the re

$$Y_{pred} = (b_1 + b_3Z)X + (b_0 + b_2(Z))$$

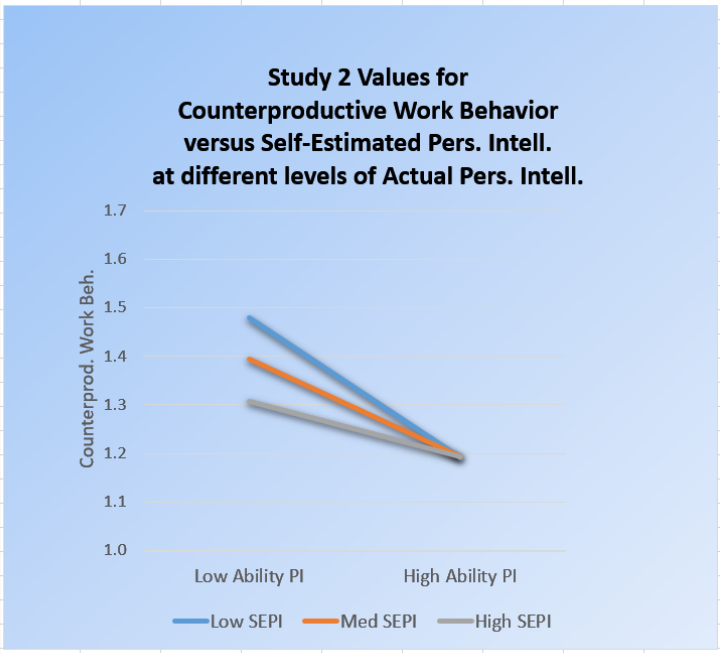
Calculate Line

$Y_{pred} = (b_1 + b_3Z)X + (b_0 + b_2(Z))$					simple slope	Predictor (Z)		Predicted Y Value	
						TOPI		TOPI	
Y predicted		X		Z		Interm. Low Val	Interm. High V	Low Value	High Value
						-1.5	1.5		
CWB	versus	TOPI	Low Z	SEPI	-0.096 times	0.144	-0.144	1.48	1.191 <-No
CWB	versus	TOPI	Medium Z	SEPI	-0.067 times	0.101	-0.101	1.39	1.1925
CWB	versus	TOPI	High Z	SEPI	-0.038 times	0.057	-0.057	1.31	1.194

Copy the values from the conclusion of the "Starting Values" page to here for your graph

	Low Ability	High Ability PI
Low SEPI	1.479	1.191
Med SEPI	1.394	1.193
High SEPI	1.308	1.194

Several Chart Options are Pr

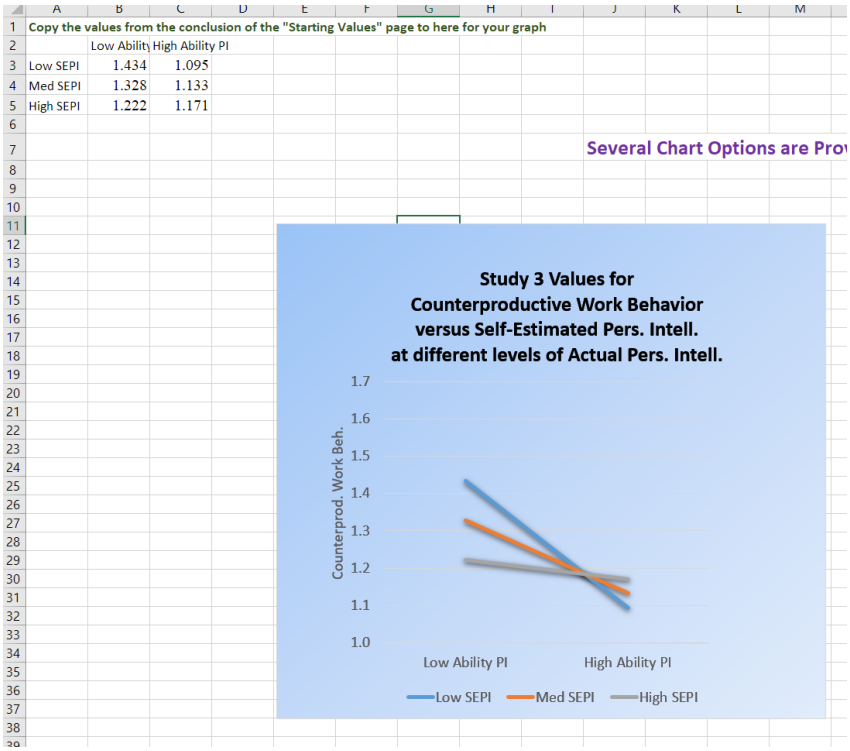


Study 3

text between tables

1	Beginning statistical setup needed for the analyses							
2	Note: Actual information and numbers to input are highlighted blue; the calculated values are highlighted yellow							
3								
4	Description of or Title for Data:	Additional analysis of Study 2 from Mayer, Lortie, Panter & Caruso, 2018						
5								
6								
7	**Step 1.** Identify the necessary variables			**Step 2.** Enter the specific values			**Step 3.** Review the calculations and read the figure from the "Plot" page.	
8	Regression Relevant Information				General Statistical Info			
9	Variable	Description	Name	Coefficient	B	Mean	Std. Dev	Coeff. Variances
10	Y	Dependent	CWB			1.2435	0.2986	0.08916196
11	Intercept	Crosses Y a	--	b_0	1.23	--	--	
12	X	predictor 1	TOPI	b_1	-0.034	0	1	1
13	Z	predictor 2	SEPI	b_2	-0.065	0	1	1
14	XZ	Interaction	iTOPISEPI	b_3	0.048	0.2795	1.03	1.0609
15								
16	Values of X desired for the interaction lines to begin and end:					-1.5	1.5	*Note:* -1.5 and 1.5 to distinguish from the Z values
17								
18	We know the mean and standard deviation of Z, so to compute the value 1 SD below, at the mean and 1 SD above the mean is fairly straightforward. The values are:							
19		Low	Medium	High				
20	Values of Z (Predictor 2)	-1	0	1	*Note:* These values are set using the M and S of Z as +/- 1 SD from M			
21	Relevant Starting Equations							
22								
23	The generalized formula for regression, with a bivariate interaction term is:							These are equations
24		$Y_{pred} = b_0 + b_1X + b_2Z + b_3(X*Z) + e$						
25								
26	These, then, are rearranged to focus on one predictor at a time, e.g., X as a predictor:							
27		$Y_{pred} = (b_1 + b_3Z)X + (b_0 + b_2(Z))$						
28								
29	The above equation has two principle terms designated as indicated below into the simple slope and simple intercept:							
30		The Simple Slope		The Simple Intercept				
31		$(b_1 + b_3Z)$		$(b_0 + b_2(Z))$				
32								
33								

34	Using the formulae above, and taking the starting values entered initially, these become											
35												
36	To find the regression line, we must set the values of Z and solve for X (we also could set X and solve for Z, but that isn't our objective for the figure)											
37	Conventionally, the values of Z are set at a low, medium, and high value, and then we solve for the Xs that will define the line relative to the predictor, Y											
38												
39		The Simple Slope		$(b_1 + b_3Z)$	Note: X is *not* part of the simple slope--that comes later							SIMPLE SLOPE
40		Z Variable	(	b_1	plus	(	b_3	times	Z	)		
41	Low Z	SEPI	(	-0.034	plus	(	0.048	times	-1)		equals	-0.113
42	Medium Z	SEPI	(	-0.034	plus	(	0.048	times	0)		equals	-0.065
43	High Z	SEPI	(	-0.034	plus	(	0.048	times	1)		equals	-0.017
44												
45		The Simple Intercept										
46		$(b_0 + b_2(Z))$										
47			(	b_0	plus	(	b_2	times	Z	)		SIMPLE INTERCEPT
48	Low Z	SEPI		1.23			-0.034		-1		equals	1.264
49	Medium Z	SEPI		1.23			-0.034		0		equals	1.230
50	High Z	SEPI		1.23			-0.034		1		equals	1.196
51												
52												
53												
54	So...using the formulae above, the simple slope and simple intercept for the three lin								Values of X (taken from row 11: "Values of X at which the regression line should begin and end"			
55			$Y_{pred} = (b_1 + b_3Z)X + (b_0 + b_2(Z))$									
56	**Calculate Line**											
57	$Y_{pred} = (b_1 + b_3Z)X + (b_0 + b_2(Z))$						simple slope		Predictor (Z)		Predicted Y Value	
58									TOPI		TOPI	
59	Y predicted		X		Z				Interm. Low Val	Interm. High V	Low Value	High Value
60									-1.5	1.5		
61	CWB	versus	TOPI	Low Z	SEPI		-0.113	times	0.170	-0.170	1.43	1.0945 **<-Now copy 3 x 2 matrix into the "Plot" page in the**
62	CWB	versus	TOPI	Medium Z	SEPI		-0.065	times	0.098	-0.098	1.33	1.1325
63	CWB	versus	TOPI	High Z	SEPI		-0.017	times	0.026	-0.026	1.22	1.1705
64												

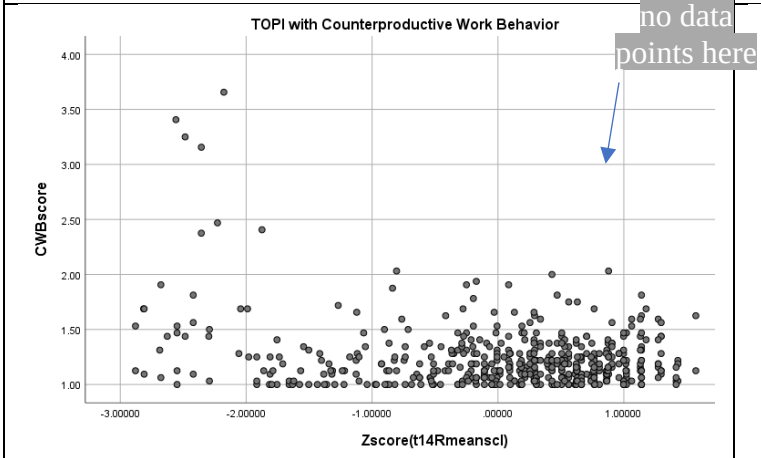
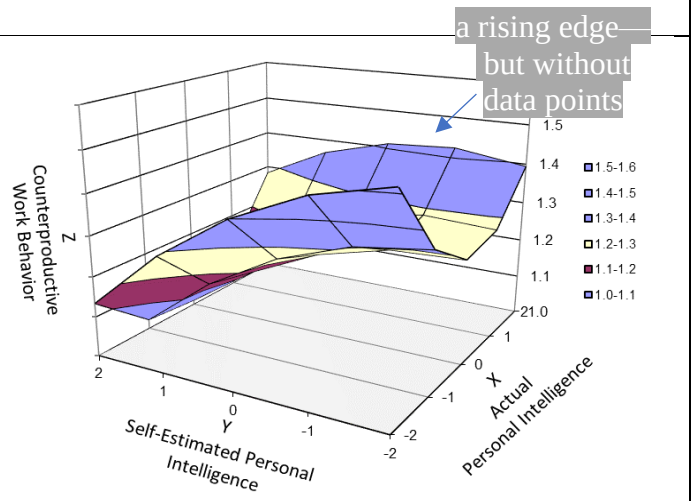


Supplement to "Response Surface Analysis for the Prediction of Counterproductive Workplace Behavior"

The prediction for counterproductive workplace behavior was far more interesting, but apparently represented a statistical artifact. It helps to revisit here Humberg et al., who observe, the RSA technique as conventionally-employed applies a parabola to fit any non-linear effects as observations diverge from their centerpoint. All modeled effects are "therefore, as a mathematical fact, symmetric around the vertical axis through its vertex…" (Humberg et. al., 2019, p. 416).

In the next figure, we show the results from Study 3. Note that the RSA to the right indicates a *reverse* congruence effect, with greater divergence between self-estimates and actual personal intelligence leading to Counterproductive Work Behavior. This curvature was statistically significant. Examining the actual scatterplot to the left, however, indicates that (presumably) the significant coefficient was a consequence of an asymmetrical effect in which only people who overestimated their personal intelligence while being quite low in their actual ability exhibited that problematic behavior (see the "no data points" note).

Appendix A. Figure 4

Self-Judged and Objective Personal Intelligence, with Counterproductive Work Behavior from Study 3	
TOPI with CWB Scatterplot	RSA "a3"—the curvilinear function, is significant
	

In fact, the marked rise in CWB among *low* TOPI scorers triggered a significant squared term in the regression (B = .05, p < .001; see Suppl. Table X.X for complete fit statistics). But because quadratic-level polynomial regression is constrained to symmetrical curves, the RSA mirrored the rise in CWB for low-TOPI individuals with a symmetrical rise for those *high* in personal intelligence. A scatterplot of the relevant data starkly contradicted the proffered model: no one high in PI also indicated high CWB. For both scatterplot and RSA. To remediate this issue would require applying spline regression: that is, curves conditional on the range of the variable-in-question (Edwards & Parry, 2017). That, in turn, complicates significance testing and violates conventions presently in use regarding the technique in the social sciences (i.e., Humberg et al., 2019).

Supplemental Table of Partial Correlations 'In addition' for Studies 1, 2 and 3

This is not mentioned in the text, but we include it here: Yet another way we thought of to consider the relative contributions of self-estimation (SEPI) and accuracy (TOPI) was (1) to partial self-estimation from the TOPI and (2) to partial ability from the SEPI. These alternative analyses are indicated for the three studies in the following supplemental tables.

Supplemental Table 5

Study 1 Correlations of Personal Intelligence Variables and Accuracy Estimates with Criterion Variables

	SEPI Estimated PI		TOPI Ability Personal Intell.		Estimated Ability.[b]	
	SEPI	SEPI *control. for ability*	TOPI-14R	TOPI-14R *control. for self-estimt.*	Absolute Accuracy	Over-Under Confidence
	Personal Intelligence – Actual, Estimated, and Accuracy-of-Estimate Variables					
SEPI	1.00					
TOPI-12	.06	.00	1.00			
Accurate Estimation	.12*	.10	.38***	.37***	1.00	
Over versus Under-Estimation[a]	.69***	1.00***	-.70***	-1.00***	-.19***	1.00
	Big Five Inventory – *Socio-Affective Styles*					
Extraversion	.39***	.38***	.05	.03	.06	.24***
Agreeableness	.34***	.34***	.09	.07	.11*	.19***
Neuroticism	-.30***	-.30***	.02	.05	-.04	-.24***
	Big Five Inventory – *Self-Control and Intellectual Styles*					
Conscientiousness	.45***	.44***	.10	.08	.15**	.25***
Openness	.20***	.20***	.09	.08	.09	.08

*p < .05, ** p < .01, *** p < .005

a. SEPI minus TOPI, high correlations are artifact of the score composite

Text between tables

Supplemental Table 6

Study 2 Correlations of Personal Intelligence Variables and Accuracy Estimates with Criterion Variables

	Ability Personal Intell.		Estimated PI		Accuracy of Est.	
	SEPI	SEPI *control. for ability*	TOPI-14R	TOPI-14R *control. for self-estimt.*	Absolute Accuracy	Over-Under Confidence
	Personal Intelligence – Actual, Estimated, and Accuracy-of-Estimate Variables					
SEPI	1.00			--		
TOPI-12	.11*	--	1.00	--	.44***	-.66***
Wordsumplus	-.05	-.10	.46***	.46***	.12*	-.37***
Absolute Accur.	.20***	.21***	.45***	.42***	1.00	.21***
Over-Under Estim.	.67***	1.00***	-.67***	-1.00***	-.20***	1.00
	Work Performance Variables					
Org. Citizen. Behav.	.18***	.17***	.03	.01	.01	.11*
Cntprdct. Wrk Beh.	-.25***	-.24***	-.19***	-.18***	-.10*	-.05
Wrk. Soc. Support	.27***	.28***	.19***	.14**	.08	.06
Job Satisfaction	.26***	.31***	-.06	-.12*	.01	.23***
Job Income	.17***	.19***	-.07	-.09	-.02	.19***

*p < .05, ** p < .01, *** p < .005

text between tables

text between tables

Supplemental Table 7

Study 3 Correlations of Personal Intelligence Variables and Accuracy Estimates with Criterion Variables

	Estimated PI		Ability Personal Intell.		Accuracy of Est.	
	SEPI	SEPI *control. for ability*	TOPI-14R	TOPI-14R *control. for self-estimt.*	Absolute Accuracy	Over-Under Confidence
	Personal Intelligence – Actual, Estimated, and Accuracy-of-Estimate Variables					
SEPI	1.00					
TOPI-14R	.28***	.00	1.00	.00		
Vocab	.10*	-.05	.49***	.48***	.09*	-.33***
Absolute Accur.	.33***	.30***	.19***	.11*	1.00	.09*
Over-Under Confid.	.60***	1.00***	-.60***	-1.00	.12**	1.00
	Big Five Inventory – *Socio-Affective Styles*					
Extraversion	.29***	.32***	-.04	-.13**	.07	.28***
Agreeableness	.37***	.34***	.18***	.08	.05	.17***
Neuroticism	-.49***	-.49***	-.03	.13**	-.22	-.38***
Core Self-Eval. Scale	.65***	.64***	.08	-.14**	.22***	.47***
	Big Five Inventory – *Self-Control and Intellectual Styles*					
Conscientiousness	.58***	.56***	.17***	.01	.19***	.34***
Openness	.35***	.31***	.20***	.11*	.05	.12**
	Employee Work Quality and Support Variables					
Org. Citizen. Beh.	.16***	.19***	-.06	.02	.01	.47***
Cntprdct. Wrk. Bh.	-.25***	-.20***	-.24***	-.19***	-.06	-.10
Soc. Support	.33***	.30***	.15***	.06	.03	.15***
Job Satisfaction	.28***	.27***	.06	-.02	.03	.18***
Job Income	.06	.07	-.03	-.05	.01	.08

*p < .05, ** p < .01, *** p < .005

Supplement to Study 1, Hypothesis 1: Self-Estimated Personal Intelligence and its Relations to Positive Affect and Self-Esteem

Before assuming that confident self-evaluations are specific to the personal intelligence domain, it is worth considering the overlap between such self-evaluations and overall positive self-evaluations on the Big Five used in Studies 1 and 3, and as a part of the Core Self Evaluation Scale used in Study 3. To determine this, we attempted to predict the SEPI from the Big Five in both studies, from the Core Self Evaluation Scale, and also added in the TOPI from both studies to see whether there might be some leftover variance that really did reflect ability at understanding personality.

The results can be seen in the following tables, which report the results from regressions from Studies 1 and 3 (no measure of the big five was included in Study 2). Note that across Studies 1 and 3, the measure of the Big Five, the measure of ability-based personal intelligence, and the sample compositions varied somewhat. Study 1 employed the BFI-44, the TOPI MINI with a sample of college students. Study 3 employed the TIPI, the TOPI 4R and online participants who were mostly middle aged and employed. Despite those differences, in both cases, the Big Five accounted for 36% and 46% of the variance of the SEPI-16 respectively. Adding in the Core Self Evaluation Scale (in a separate analysis not in the table) accounted for another 7% of the variance, or about 53% overall. The leftover variance unique to the TOPI was essentially zero in Study 1 and 3% on Study 3, indicating that responses to the SEPI reflected confidence more than any actual ability at a ratio of something between the entirety of variance for Study 1 to a ratio of confidence-to-actual-ability of 20 to 1 in Study 3. In other words, both studies point to the fact that the SEPI is near-completely a measure of confidence.

text between tables

Supplemental Table 8 for Study 1

Predicting Self-Estimated Personal Intelligence (SEPI) from the Big Five and TOPI MINI (*N* = 346)

Steps	**(1) The Big Five**			**(2) The Big Five, CSES, and Personal Intelligence**		
	B (SE of Coef.)	95% conf.	*β*	B	95% conf. int.	*β*
Extraversion	.313 (.061)	.193 to .433	.24	.313	.240 to .434	.24
Agreeableness	.229 .082	.067 to .390	.13	.229	.132 to .391	.13
Neuroticism	-.228 .066	-.358 to -.097	-.16	-.227	-.159 to -.097	-.16
Conscientious.	.610 .081	.451 to .770	.35	.611	.348 to .771	.35
Openness	.197 .082	.036 to .358	.11	.197	.106 to .359	.11

TOPI MINI	--			-.015	-.002 to .509	-.00
	Summary of Regression Model			Summary of Regression Model		
R	.602***	--	--	.602***	--	--
R-sqr.	.363***	--	--	.363***	--	--
Adj. R-sqr.	.353***	--	--	.351***	--	--
Sign. of Change	$F_{change}(5,341) = 38.81$***			$F_{change}(1, 340) = .003$		

p < .05, **p < .10, ***p < .001
All regressions include a constant term but we omit these coefficients for brevity.

text between table

text between table

Supplemental Table 9 for Study 3

Predicting Self-Estimated Personal Intelligence (SEPI) from the Big Five and TOPI MINI (N = 346)

Steps	**(1) The Big Five**			**(2) The Big Five, CSES, and Personal Intelligence**		
	B (SE of Coef.)	95% conf.	β	B	95% conf. int.	β
Extraversion	0.045 (.014)	0.019 to 0.072	0.12	0.032 (.013)	0.007 to 0.057	0.08
Agreeableness	0.065 (.019)	0.028 to 0.103	0.13	0.039 (.017)	0.005 to 0.074	0.08
Neuroticism	-0.089 (.018)	-0.124 to -.054	-0.20	-0.026 (.018)	-0.061 to 0.010	-0.06
Conscientious.	0.211 (.021)	0.169 to 0.253	0.39	0.143 (.020)	0.103 to 0.183	0.26
Openness	0.077 (.018)	0.042 to 0.112	0.16	0.046 (.017)	0.014 to 0.079	0.09
CSES	--	--		0.347 (.039)	0.270 to 0.423	0.39
TOPI MINI	--	--		0.010 (.002)	0.007 to 0.014	0.18
	Summary of Regression Model			Summary of Regression Model		
R	.678***	--		.749***	--	--
R-sqr.	.459***	--		.561***	--	--
Adj. R-sqr.	.453***	--		.555***	--	--
Sign. of Change	$F_{change}(5,475) = 80.66$***			$F_{change}(2,473) = 55.11$***		

p < .05, **p < .10, ***p < .001
All regressions include a constant term but we omit these coefficients for brevity.

text between tables

References

DeCoster, J., Gallucci, M., & Iselin, A.-M. R. (2011). Best practices for using median splits, artificial categorization, and their continuous alternatives. *Journal of Experimental Psychopathology*, *2*(2), 197–209. psyh. https://doi.org/10.5127/jep.008310

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis.* (2001-05785-010; pp. 350–400). Jossey-Bass; psyh. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-05785-010&site=ehost-live

Edwards, J. R. (2009). Latent variable modeling in congruence research: Current problems and future directions. *Organizational Research Methods*, *12*(1), 34–62. psyh. https://doi.org/10.1177/1094428107308920

He, J. C., & Côté, S. (2019). Self-insight into emotional and cognitive abilities is not related to higher adjustment. *Nature Human Behavior*, *3*, 867–884. https://doi.org/10.1038/s41562-019-0644-0

Humberg, S., Nestler, S., & Back, M. D. (2019). Response surface analysis in personality and social psychology: Checklist and clarifications for the case of congruence hypotheses. *Social Psychological and Personality Science*, *10*(3), 409–419. psyh. https://doi.org/10.1177/1948550618757600

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, *66*(1), 206–219. psyh. https://doi.org/10.1037/0022-3514.66.1.206

Krueger, J. I., Heck, P. R., & Asendorpf, J. B. (2017). Self-enhancement: Conceptualization and assessment. *Collabra Psychology*, *3*(28). http://dx.doi.org/10.1525/collabra.91

Laird, R. D., & Weems, C. F. (2011). The equivalence of regression models using difference scores and models using separate scores for each informant: Implications for the study of informant discrepancies. *Psychological Assessment*, *23*(2), 388–397. psyh. https://doi.org/10.1037/a0021926

Linn, R. L., & Slinde, J. A. (1977). The Determination of the Significance of Change between Pre- and Posttesting Periods. *Review of Educational Research*, *47*(1), 31.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. psyh. https://doi.org/10.1037/1082-989X.7.1.19

Mayer, J. D., Caruso, D. R., & Panter, A. T. (2019). Advancing the measurement of personal intelligence with the Test of Personal Intelligence, Version 5 (TOPI 5). *Journal of Intelligence*, *7*(4), 1–17. https://doi.org/doi.org/10.3390/jintelligence7010004

Mayer, J. D., Lortie, B., Panter, A. T., & Caruso, D. R. (2018). Employees high in personal intelligence differ from their colleagues in workplace perceptions and behavior. *Journal of Personality Assessment*, *100*(5), 539–550. psyh. https://doi.org/10.1080/00223891.2018.1455690

Mayer, J. D., Panter, A. T., & Caruso, D. R. (n.d.). *Alternative measures of personal intelliigence* [Unpublished raw data].

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2012). Does personal intelligence exist? Evidence from a new ability-based measure. *Journal of Personality Assessment*, *94*(Journal Article), 124–140. https://doi.org/10.1080/00223891.2011.646108

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2017). A closer look at the Test of Personal Intelligence (TOPI). *Personality and Individual Differences*, *111*(Journal Article), 301–311. https://doi.org/10.1016/j.paid.2017.02.008

Mayer, J. D., & Skimmyhorn, W. (2017). Personality attributes that predict cadet performance at West Point. *Journal of Research in Personality*, *66*, 14–26. psyh. https://doi.org/10.1016/j.jrp.2016.10.012

Schönbrodt, F. D., Humberg, S., & Nestler, S. (2018). Testing similarity effects with dyadic response surface analysis. *European Journal of Personality*. psyh. https://doi.org/10.1002/per.2169

Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming

limitations of difference scores. *Journal of Business and Psychology, 25*(4), 543–554. psyh. https://doi.org/10.1007/s10869-010-9183-4

Trafimow, D. (2015). A defense against the alleged unreliability of difference scores. *Cogent Mathematics*, *2*(1), 1–11. a2h.

Appendix A. Polynomial Regressions Used for the RSAs

Appendix A, Table 1 (next page) contains the polynomial regressions that were used for the response surface analyses earlier in this document.

buffer text

Appendix A. Table 1

Statistics for the Polynomial Regression and Response Surface Analyses for Conscientiousness, Vocabulary, and Counterproductive Work Behavior

			Conscientiousness				**Vocabulary**				**Counterproductive Work Behavior**			
			Study 1		Study 3		Study 2		Study 3		Study 2		Study 3	
Variable Name	Regression Term	Weight	unstand. beta weights	standard error	unstand. beta weights	standard error	unstand. beta weights	standard error	unstand. beta weights	standard error	unstand. beta weights	standard error	unstand. beta weights	standard error
	constant	B_o	3.585	.036	5.721	.067	.757	.010	.722	.011	1.281		1.208	
TOPI	x	b_1	-.021	.031	-.058	.059	.082***	.012	.108***	.01	-0.001	.021	.003	.017
SEPI	y	b_2	.255***	.28	.695***	.049	-.015*[b]	.008	-.007	.008	-0.071***	.015	-.062***	.014
$TOPI^2$	x^2	b_3	-.027*[b]	.014	-.037	.043	0	.004	.017*	.007	.02**	.008	.049***	.012
TOPI x SEPI	xy	b_4	.037	.026	-.065	.054	-.002	.007	.003	.006	.019	.013	.017	.016
$SEPI^2$	y^2	b_5	.011	.019	.017	.035	.004	.005	-.002	.009	-.007	.009	-.018	.01
Cov(X,Y)	--	b1b2	1.67E-05		-1.00E-03		-4.15E-06		-1.34E-05		9.23E-06		-1.75E-05	
Cov(X^2,XY)	--	b3b4	0.00E+00		-1.00E-03		-8.51E-05		-2.97E-05		2.87E-05		-9.56E-05	
Cov(X2,Y2)	--	b3b5	-1.03E-05		-5.19E-05		-4.15E-07		-1.34E-06		-1.50E-06		-4.35E-06	
Cov(XY,Y2)	--	b4b5	-8.90E-.05		-7.20E-05		1.34E-06		-1.86E-06		4.48E-06		-5.90E-06	
Slopes and Curves[a]			coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
slope along x = y	a_1		.23***	.04	.64***	.06	.07***	.01	.10***	.01	-.07**	.03	-.06**	.02
curvature on X = y	a_2		.02	.03	-.09	.06	.00	.00	.02	.01	.03	.02	.05***	.02
slope along x = -y	a_3		-.28***	.04	-.75***	.09	.10***	.02	.12***	.01	.07**	.03	.07***	.02
curvature on x = -y	a_4		-.05	.04	.05	.09	.01	.02	.01	.02	-.01	.02	.01	.03

a. Studies are as described in the paper: Study 1 is Alternate Uses data, Studies 2 and 3 are data collected by Moore and Lortie, respectively and described elsewhere (Mayer et al., 2018, Studies 1 and 2).

b. X and Y have been converted to z-scores; unstandardized *b* coefficients are used

The polynomial regression equation is $Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2$ (Humberg et al., 2019, Equation 1)

The line of congruence (LOC) is $Z = b_o + a_1X + a_2X^2$. For the line of incongruence (LOIC), $Y = -X$ in the polynomial above, is $Z = b_o + a_3X + a_4X^2$

buffer text

section break here

Discussion of the Results of the Response Surface Analysis

All-in-all, the results from the RSA were mixed. On the one hand, the results very strongly supported our use of DIZs. There were no signs of overcorrection. The RSA results also were consistent with the oft-noted issue of under-correction of residual scores (Appendix A, Figure 1). These both were welcome findings in that they provided support for the use of such scores and for our understanding of them.

On the other hand, the RSAs failed to detect the same congruence effects as the DIZs had detected with consistency across studies—and as we obtained in full linear regressions for conscientiousness, vocabulary, and counterproductive work behavior. There are several possible explanations of the polynomial-with-RSA's failures to find effects.

For one, RSA often seems to fail to find effects. For example, in a recent pre-registered study, He and Côté (2019) failed to find expected effects of nearly every kind including convergence between self-estimates and actual intelligence, despite using roughly a thousand participants and, in their pre-registered study, testing for a substantial number of hypotheses—and cited similarly disappointing uses of the technique. Arguing against that, of course, are the lovely convergence effects we obtained for the DIZs.

A second possibility was that perhaps there were violations of various assumptions in the data, ranging from asymmetrical curvature to skew to other departures from assumptions that impeded the technique's capacity to detect effects.

A third possibility is that the DIZ scores, based as they are on specific individual responding rather than the idealized group-level relationships of any kind of regression, retain some key information at the level of an individual's estimation of their reasoning level that is both reliable and valid, and simply does not readily conform to the somewhat idealized group

lines and curves entailed by best-fit models. It could be this individual variability, a variability that violates easy prediction, that provides the key to examining what correlates of accuracy truly exist.

Neither did we overlook a further, fourth, possibility that the effects simply are not there. As we indicate in the discussion, we hope to follow-up on this work with further research and further replication beyond the replications here. And yet, the results do seem surprisingly interpretable, meaningful, and to arise using DIZs, residuals, and not only in simple correlations, but also in regressions (see the section on regressions).

Looking over these possibilities, we acknowledge we simply do not understand the discrepancy. That said, we do have some confidence that the DIZs and, to a lesser extent, the residual scores both may provide fruitful for future research.

Section Break Here

Appendix B: The "High-Low" Four Group Approach

Aims of the Four-Group Approach

In the following studies, we examine people whose personal intelligence and *beliefs about* their personal intelligence agree or disagree. To do so we examined the relation between those abilities and beliefs in two ways. In our primary approach, we divided people into four groups consisting of people who were:

(a) low in personal intelligence and knew it (low-low),

(b) high in personal intelligence and knew it (high-high),

(c) high in personal intelligence, but who lacked confidence in their abilities (high-low) and

(d) low in personal intelligence, but were overconfident in their abilities (low-high).

The four groups provide an important perspective on the nature of patterns of true ability and self-confidence. In addition, we created two scores for each person representing the accuracy of their self-estimated personal intelligence, i.e., their estimate's proximity to their true ability, and each person's over- (or under-) confidence in estimating their ability.

To understand the nature of the groups and their scores, we examined the groups in relation to their average scores on socio-affective traits such as extraversion and traits of self-control such as conscientiousness. We also examine their verbal intelligence, and lifespace (e.g., act-frequency) data concerning their behavior at work. We also conducted correlational explorations of people's accuracy and under-over-confidence scores. In so doing, we develop a preliminary picture of the similarities and differences among people with varied levels of personal intelligence, who have matched or mismatched levels of confidence.

Common Rules and Procedures Applied to Each Sample

Across the three data sets we applied the following common procedures:

- First, we used the data set as it was employed in any published report: That is, we followed the original method of screening and employed the unaltered data set.

- Second, we used the SEPI-16 version of the Self-Estimated Personal Intelligence scale, as that is highly reliable, brief, and common across all three studies and reported in the published studies. (The SEPI versions used in the Alternative Uses dataset and Study 1 of Mayer et al., 2018 had 120 and 74 items respectively. Both however, included all 16 items of the SEPI-16).
- Third, in creating the four high-high, high-low, etc., ability versus estimate groups described in the Introduction and other parts of the document, we always used within-sample statistics to split groups at the median of the TOPI (whichever form) and the SEPI-16

The groups are somewhat uneven in size for two reasons. First, the TOPI-MINI has only 12 intervals and so the "median" split was often five or more percentiles away from an even division. Second, the TOPI and SEPI tend to exhibit a positive correlation with one another, which slightly increments the size of the Low-Low and High-High groups relative to the others. That said, all four group sizes in the three studies are reasonable in size. The median splits are indicated in Appendix A, Table 1.

text between tables

Appendix Table B1

Cutpoints Nearest the 50th Percentile (Median) of the Distribution for the TOPI and SEPI

Measure	Study 1		Study 2		Study 3	
	less than	greater or equal than	less than	greater or equal than	less than	greater or equal than
TOPI MINI	.84[a]	.84	.84[a]	.84	--	--
TOPI 14R	--	--	--	--	52*	52
SEPI 16	3.65	3.65	3.71	3.71	3.94	3.94

a. The .84 proportion correct cutpoints for the TOPI MINI basically are saying that test-takers either scored 10 or fewer (a proportion of .83) or 11 or more correct.

*The TOPI 14R uses scaled scores approximating a T-Scale (M = 50, S = 10). Backward lookup indicates the score of 52 is approximately equal to a proportion of .74 correct. The discrepancy between the TOPI MINI .83 (see "a" above) and the .74 of the TOPI 14R is due to the easier quality (i.e., higher proportional score levels) of the TOPI MINI relative to the full TOPI 14R.

text between tables

Appendix B. Study 1 Using the Four Group Approach

text divider between tables

Appendix Table B2

Descriptive Statistics for the Key Measures of Study 1

	Overall		Group 1	Group 2	Group 3	Group 4	
	Mean	Standard Deviation	Low TOPI/ Low SEPI	High TOPI / Low SEPI	Low TOPI / High SEPI	High TOPI / High SEPI	
	Group Demographics, Personal Intelligence, Self-Estimates, and Related Scores						
	Breakdown by Gender and Age						Statistical Test
Gender [c]	Men	97	36	21	23	17	$\chi^2(3) = 9.23$*
	Women	250	54	65	67	64	
proportion women	.72		.60	.76	.74	.79	
Age	19.70	2.32	19.53	19.77	19.49	20.03	$F(3,338) = .921^{ns}$
	N	347	90	86	90	81	--
	Accuracy and Confidence Scores (on T-Scales)						
Absolute Accur.	50.0	10.0	51.7	45.3	46.9	56.6	--
Confidence	50.0	10.0	50.1	39.3	59.8	50.4	
	Actual and Estimated Personal Intelligence						Reliabilities
TOPI-MINI-12	.82	.45	.69	.95	.74	.94	.64
SEPI-16	3.63	.16	3.36	3.39	3.76	3.77	.83
	Criterion Test Means by Group						
	Big Five Inventory – *Socio-Affective Styles*						
Extraversion	3.44	.77	3.16	3.29	3.64	3.67	.87
Agreeableness	3.81	.58	3.65	3.66	3.99	3.81	.77
Neuroticism	3.13	.71	3.32	3.26	2.94	3.01	.81
	Big Five Inventory – *Self-Control and Intellectual Styles*						
Conscientiousness	3.57	.57	3.40	3.31	3.77	3.81	.78
Openness	3.51	.54	3.44	3.42	3.53	3.65	.75

^{ns}not significant; *$p < 05$; **$p < .01$; **$p < .005$

a. cutpoints for sample divisions (nearest the median) were "lower than" and "higher or equal to" 0.9^a and 3.65^a

b. Estimated accuracy was calculated as follows: First, the difference between the test-taker's estimated ability (SEPI) and their actual ability (TOPI) was calculated by first converting scores on each measure to z-scores (i.e., unit standardized values) and then subtracting one from the other. Next, that value was converted to an absolute value (directionality was discarded) multiplied by negative 1.0 to obtain an accuracy rather than a discrepancy score. Finally, the accuracy score was itself converted to a T-score with $M = 50$ and $S = 10$.

x. Three participants indicated they were non-gender-binary and are not included in the analysis for gender but are included in other analyses

c. There were more men in some groups than expected, $\chi^2(3) = 9.23$*, $p < .05$

d. There was no difference across groups in age, $F(3,338) = .921^{ns}$

text divider between tables

text divider between tables

Appendix Table B3

Study 1 Comparisons on the Big Five Between Groups High and Low in Self-Estimated Personal Intelligence, in Actual Personal Intelligence, and in High and Low Accuracy of Self-Estimates

	High v. Low SEPI Group Means			High v. Low TOPI Group Means			High v. Low Accuracy Group Means[a]		
	Low SEPI	High SEPI	sig. diff. t-test	Low TOPI	High TOPI	sig. diff t-test	Low-Accuracy	High-Accuracy	sig. diff. t-test
N	176	171	--	180	167	--	90	81	--
	Accuracy and Confidence Scales								
Accuracy									
Confidence									
	Big Five (BFI-44) ***Socio-Affective Styles***								
Extraversion	3.22	3.65	-5.39***	3.39	3.48	-.92^{ns}	3.64	3.67	.78^{ns}
Agreeableness	3.65	3.97	-5.57***	3.82	3.80	.30^{ns}	3.99	3.81	.58^{ns}
Neuroticism	3.29	2.97	4.29***	3.13	3.14	-.13^{ns}	2.94	3.01	-.92^{ns}
	Big Five (BFI-44) ***Self-Control and Intellectual Styles***								
Conscientious.	3.36	3.79	-7.58***	3.59	3.55	.58^{ns}	3.77	3.81	-.74^{ns}
Openness	3.43	3.56	-2.73**	3.49	3.53	-.76^{ns}	3.53	3.65	-1.0^{ns}

[a]The Low-Low and High-High Groups (1 and 4) versus Low-High and High-Low (2 and 3)

text divider between tables

Appendix Table B3 (Alternate)

Study 1, Alternative Table 3 (same information, different format)

Study 1 Comparisons on the Big Five Between Groups High and Low in Self-Estimated Personal Intelligence, in Actual Personal Intelligence, and in High and Low Accuracy of Self-Estimates

		Traits of Socio-Affective Styles			***Traits of Self-Control and Intellect. Styles***	
	N	Extraversion	Agreeableness	Neuroticism	Conscientious.	Openness
	Comparison of Groups High and Low on Self-Estimated Personal Intelligence					
Low SEPI	176	3.22	3.65	3.29	3.36	3.43
High SEPI	171	3.65	3.97	2.97	3.79	3.56
t-test for difference		-5.39***	-5.57***	4.29***	-7.58***	-2.73**
	Comparison of Groups High and Low on Actual, Ability-Based Personal Intelligence					
Low TOPI	180	3.39	3.82	3.13	3.59	3.49
High TOPI	167	3.48	3.80	3.14	3.55	3.53
t-value for sig. diff		-.92^{ns}	.30^{ns}	-.13^{ns}	.58^{ns}	-.76^{ns}
	Comparison of Groups of High and Low Accuracy[a]					
Low-Accuracy Groups	176	3.64	3.99	2.94	3.77	3.53
High-Accuracy Groups	171	3.67	3.81	3.01	3.81	3.65
t-value for sig. diff		.78^{ns}	.58^{ns}	-.92^{ns}	-.74^{ns}	-1.0^{ns}

[a]The Low-Low and High-High Groups (1 and 4) versus Low-High and High-Low (2 and 3)

text divider between tables

Test of Hypotheses

Was the correlation between self-estimated and actual personal intelligence between *r* = .20 and .30 (Hypothesis 1)? The correlation between the SEPI16 and MINI-12 in this sample was *r* = .06, n.s., somewhat lower than expected (and usually obtained).

Did people higher in personal intelligence have more accurate self-estimates of their abilities than people lower in personal intelligence, as evaluated by a regression using the absolute discrepancy of the TOPI-MINI and SEPI-16 scores in standardized form.) An accuracy score was constructed equal to -1 times the discrepancy between the z-scores of the TOPI MINI and SEPI 16. In a stepwise regression predicting accuracy, the TOPI MINI predicted accuracy with a standardized *b* = .37, *p* < .000. We also tried to predict accuracy with the SEPI-16, yielding a *b* = .095, n.s..

text divider between tables

Study 2 Using the Four Group Approach

The characteristics of the overall sample and subgroups are indicated in Study 2, Table 1.

Appendix Table B4

Study 2: Breakdown of Age and Gender by Group

		Overall	Group 1	Group 2	Group 3	Group 4	Statistical Test[a]
		Mean (Std. Dev.)	Low TOPI/ Low SEPI	High TOPI / Low SEPI	Low TOPI / High SEPI	High TOPI / High SEPI	
Gender [b]	Men	176	41	47	25	63	$\chi^2(3) = 4.20^{ns}$
	Women	217	41	70	41	65	
proportion women		.55	.50	.60	.61	.51	
Age M(S)		19.70 (2.32)	19.53 (1.82)	19.77 (1.68)	19.49 (1.60)	20.03 (3.71)	$F(3,390) = 1.12^{ns}$
N[b]		393	82	117	67	128	--

[ns]not significant; *p < 05; **p < .01; **p < .005

a. A Pearson chi-square for gender; a one-way ANOVA for age

b. One participant's data was missing for gender

text between tables

Test of Hypotheses

Was the correlation between self-estimated and actual personal intelligence between *r* = .20 and .30 (Hypothesis 1)? The correlation between the SEPI16 and MINI-12 in this sample was $r = .11$, $p < .03$, somewhat lower than expected (and usually obtained), but higher than in Study 1.

Did people who judged themselves as high in personal intelligence on the SEPI exhibit more positive self-judgments in general, as indicated by higher scores on the Job Satisfaction, Organizational Citizenship Scales, and by lower scores on the Counterproductive Work Behavior scale (Hypothesis 2)? The relevant comparison is indicated in Study 2, Table 2. As indicated there, there were considerable significant differences between the high-low SEPI groups on all the measures: not only those hypothesized above, but also on perceived social support at work. The high SEPI scorers basically reported that they were better behaved (from a social desirability standpoint) and better off on all four characteristics—more helpful, less destructive, better socially supported, and more satisfied overall at work.

text between tables

Appendix Table B5

Study 2: Descriptive Statistics for the Key Measures,

	Overall		Group 1	Group 2	Group 3	Group 4	Reliability
	Mean	Standard Deviation	Low TOPI/ Low SEPI	High TOPI / Low SEPI	Low TOPI / High SEPI	High TOPI / High SEPI	
	Personal and Verbal Intelligences, and Self-Estimates of Personal Intelligence						
WordSumPlus	.76	.18	.68	.84	.69	.78	.73
TOPI-MINI-12	.87	.15	.72	.95	.76	.96	.71
SEPI-16	3.7	.56	3.2	3.3	4.1	4.1	.89
Absolute Accur.	50.00	10.00	51.78	46.58	42.44	55.96	--

Accuracy w. Direct.	50.00	10.00	52.13	40.37	62.47	50.93	--
	Work Quality and Support Variables						
Org. Citizen. Behav.	2.9	.68	2.7	2.8	3.1	2.9	.92
Cntprdct. Wrk Beh.	1.2	.21	1.4	1.4	1.2	1.2	.90
Wrk. Soc. Support	3.8	.65	3.5	3.7	4.0	4.0	.92
Job Satisfaction	5.3	1.5	4.9	5.0	5.9	5.4	.79
N	394		82	117	67	128	--

text between tables

Did people higher in personal intelligence have more accurate self-estimates of their abilities than people lower in personal intelligence, as evaluated by a regression using the absolute discrepancy of the TOPI-MINI and SEPI-16 scores in standardized form.) The accuracy score was constructed as in Study 1. In a stepwise regression predicting accuracy the TOPI MINI again predicted accuracy with a standardized $b = .436$, $p < .001$, and this time, the SEPI was slight more predictive of accuracy than its marginal level before $b = .153$, $p < .001$. Details are in Study 2, Table 3.

Were there differences on other traits between those who more accurately predicted their personal intelligence (the High-High and Low-Low groups) versus those with discrepant estimates of personal intelligence (the HI-LO and LO-HI) groups (Hypothesis 4)? As indicated in the right-most columns of Study 2, Table 4, none of the alternative work-centered traits predicted accuracy in understanding personality in the form of TOPI-MINI scores.

text divider between tables

Appendix Table B6

Study 2 Comparisons on the Big Five Between Groups High and Low in Self-Estimated Personal Intelligence, in Actual Personal Intelligence, and in High and Low Accuracy of Self-Estimates

	High v. Low SEPI Group Means			High v. Low TOPI Group Means			High v. Low Accuracy Group Means		
	Low SEPI	High SEPI	sig. diff. t-test	Low TOPI	High TOPI	sig. diff t-test	Low-Accuracy	High-Accuracy	sig. diff. t-test
	Employee Work Quality and Support Variables								
Org. Citizen. Beh.	2.77	2.98	3.13***	2.89	2.86	-0.47	2.90	2.86	$.54^{ns}$

Cntprdct. Wrk Bh.	1.36	1.23	-4.38****	1.32	1.28	-1.34	1.30	1.29	.431[ns]
Soc. Support	3.65	3.99	5.44****	3.74	3.86	-1.67	3.85	3.79	.768 [ns]
Job Satisfaction	4.94	5.60	4.57****	5.35	5.21	-0.88	5.29	5.24	.361 [ns]
N	198	195	--	149	245	--	90	81	--

*p < .05, ** p < .01, *** p < .005, p < .001

text divider between tables

Appendix B. Study 3 Using the Four-Group Approach

text divider between tables

Appendix Table B7

Study 3: Breakdown of Age and Gender by Group

		Overall	Group 1	Group 2	Group 3	Group 4	Statistical Test[a]
		Mean (Std. Dev.)	Low TOPI/ Low SEPI	High TOPI / Low SEPI	Low TOPI / High SEPI	High TOPI / High SEPI	
Gender [b]	Men	230	93	44	36	57	$\chi^2(3) = 21.06^{***}$
	Women	250	53	61	55	81	
proportion women		.52	.36	.58	.60	.59	
Age M(S)		19.70 (2.32)	19.53 (1.82)	19.77 (1.68)	19.49 (1.60)	20.03 (3.71)	$F(3,478) = 4.66^{**}$
N[b]		480	146	105	91	138	--

[ns]not significant; *p < 05; **p < .01; ***p < .005

a. A Pearson chi-square for gender; a one-way ANOVA for age

b. Two participants' data was missing for gender

text divider between tables

text divider between tables

Appendix Table B8

Study 3: Descriptive Statistics for the Key Measures,

	Overall		Group 1	Group 2	Group 3	Group 4	Reliability
	Mean	Standard Deviation	Low TOPI/ Low SEPI	High TOPI / Low SEPI	Low TOPI / High SEPI	High TOPI / High SEPI	
	Personal and Verbal Intelligences, and Self-Estimates of Personal Intelligence						
WordSumPlus	10.35	2.49	9.24	11.46	9.27	11.35	.73
TOPI 1.4R	49.23	10.90	38.51	56.95	44.49	57.70	.94
SEPI-16	3.87	.64	3.35	3.42	4.37	4.38	.95
	Accuracy Variables (on a T-Scale)						
Absolute Accur.	50.00	10.00	50.82	43.63	46.30	56.51	--
Over-Under Conf.	50.00	10.00	51.48	38.19	60.18	50.87	--
	Big Five Inventory and Related – *Socio-Affective Styles*						
Extraversion	3.93	1.65	3.65	3.43	4.64	4.16	.76
Agreeableness	5.06	1.19	5.26	6.08	5.94	5.55	.58
Neuroticism	2.79	1.45	3.22	3.49	2.00	2.31	.74
Core Self-Eval. Scl	3.73	.70	3.40	3.32	4.12	3.73	.90
	Big Five Inventory– *Self-Control and Intellectual Styles*						
Conscientiousness	5.68	1.18	5.12	5.15	6.31	6.28	.63
Openness	5.17	1.29	4.69	4.90	5.52	5.66	.55
	Work Quality and Support Variables						
Org. Citizen. Behav.	2.92	.66	2.92	2.76	2.99	2.99	.92
Cntprdct. Wrk Beh.	1.24	.30	1.34	1.24	1.17	1.24	.93
Wrk. Soc. Support	3.90	.63	3.68	3.83	4.07	4.09	.80
Job Satisfaction	5.31	1.55	4.98	5.02	5.49	5.76	.94
N	481		145	107	91	138	--

text divider between tables

Test of Hypotheses

Was the correlation between self-estimated and actual personal intelligence between *r* = .20 and .30 (Hypothesis 1)? The correlation between the SEPI16 and MINI-12 in this sample was $r = .28$, $p < .001$, somewhat higher than typically obtained. The hypothesis was supported.

Did people who judged themselves as high in personal intelligence on the SEPI exhibit more positive self-judgments in general, as indicated by higher scores on the Positive Affect scores of the Big Five, as well as on Job Satisfaction, Organizational Citizenship Scales, and by lower scores on the Counterproductive Work Behavior scale (Hypothesis 2)? The relevant comparison is indicated in Study 3, Table 3. As indicated there, there were considerable significant differences between the high-low SEPI groups on all the measures: The high SEPI scorers basically reported on the Big Five reported that they were livelier, more sociable, agreeable, less anxious, more conscientious and open, as well as (more generally) better behaved, less destructive, better socially supported, and more satisfied overall at work. The high SEPI scorers are, in other words, a happier, more positive and satisfied group.

Did people higher in personal intelligence have more accurate self-estimates of their abilities than people lower in personal intelligence? Once again, we asked the question using absolute discrepancy of the standardized TOPI 1.4R and SEPI-16 scores. In the same stepwise regression as earlier studies, the TOPI 1.4R again predicted accuracy with a standardized $b = .10$, $p < .05$, and this time, the SEPI was more predictive of accuracy than the TOPI $b = .31$, $p < .001$. In a regression by itself, the TOPI 1.4R predicted at a $b = .19$, closer to the earlier studies. It seems likely that the higher correlation between TOPI and SEPI accounted for the somewhat weaker prediction in the initial regression.

Were there differences on other traits between those who more accurately predicted their personal intelligence (the High-High and Low-Low groups) versus those with discrepant estimates of personal intelligence (the HI-LO and LO-HI) groups (Hypothesis 4)? As indicated in the right-most columns of Table 8.2, of the ten variables tested, nine were non-signficant distinguishers, following the pattern of earlier studies. Only the

Counterproductive Work Behavior score appeared relevant: Participants none of the alternative work-centered traits predicted accuracy in understanding personality in the form of TOPI-MINI scores.

text divider between tables

Appendix Table B9

Study 3. Comparisons on the Big Five Between Groups High and Low in Self-Estimated Personal Intelligence, in Actual Personal Intelligence, and in High and Low Accuracy of Self-Estimates

	High v. Low SEPI Group Means			High v. Low TOPI Group Means			High v. Low Accuracy Group Means		
	Low SEPI	High SEPI	sig. diff. t-test	Low TOPI	High TOPI	sig. diff t-test	Low-Accuracy	High-Accuracy	sig. diff. t-test
	Big Five (BFI-44) and Related – *Socio-Affective Styles*								
Extraversion	1.44	1.76	-5.41***	4.03	3.84	1.30	3.99	3.90	$.59^{ns}$
Agreeableness	5.14	6.00	-7.94***	5.45	5.65	-1.70	5.64	5.49	1.30^{ns}
Neuroticism	3.33	2.19	9.43***	2.75	2.82	-.58	2.80	2.78	$.20^{ns}$
CSES	3.37	4.14	-14.43***	3.70	3.77	-1.15	3.71	3.75	$-.70^{ns}$
	Big Five (BFI-44) – *Self-Control and Intellectual Styles*								
Conscientious.	5.13	6.29	-12.34***	5.58	5.79	-1.97*	5.68	5.69	$-.03^{ns}$
Openness	4.78	5.61	-7.41***	5.01	5.32	-2.70**	5.18	5.17	$.15^{ns}$
	Employee Work Quality and Support Variables								
Org. Citizen. Beh.	2.86	2.99	-2.14*	2.95	2.89	1.03	2.87	2.96	-1.48^{ns}
Cntprdct. Wrk Bh.	1.30	1.18	4.37***	1.27	1.21	2.19*	1.21	1.27	-2.10*
Soc. Support	3.74	4.08	-6.10***	3.83	3.98	-2.62**	3.94	3.87	1.15^{ns}
Job Satisfaction	5.00	5.66	-4.75***	5.18	5.44	-1.85	5.24	5.36	$-.836^{ns}$
N									

*p < .05, ** p < .01, *** p < .005, p < .001

a. cutpoints for sample divisions (nearest the median) were “lower than” and “higher or equal to” 0.9^{a} and 3.65^{a}

text divider between tables

Appendix C. A Regression Approach

Comparative Regression Analyses

We also conducted comparative multiple regressions of several sets of scores to predict (a) socio-affective traits, (b) traits of intelligence, intellectual style, and self-control, and (c) workplace attitudes and behaviors. We compared three approaches to our measures. Model 1 employed the original SEPI and TOPI scales, and their interactions. Model 2 employed DIZ scores for confidence and accuracy and their interactions, and Model 3 was the same except that the regression scores were used. The comparative results can be seen in Appendix C Table 1.

Appendix C Table 1

Comparison Regressions on Criterion Variables Using Three Models, Each Employing Original SEPI and TOPI or Calculated Scores of Accuracy and Overconfidence

		Model 1 Original SEPI and TOPI Scores					Model 2 DIZ Scores for Confidence and Accuracy				Model 3 Regression Scores for Confidence and Accuracy			
Variable	Std.	Mult. R	SEPI[b]	TOPI[b]	Interact		Mult. R	OvrConf	Acc	Inter.	Mult. R	OvrConf	Acc	Interact
		Socio-affective traits												
Extrav.	1	.389***	.389***	.039	.046		.265***	.273***	.104[a]	.013	.392***	.389***	-.001	.084[a]
	3	.317***	.327***	-.130**	.005		.278***	.279***	.036	.011	.315***	.319***	-.007	.016
Agree.	1	.354***	.346***	.080	.062		.244***	.285***	.127*	.098	.368***	.328***	.135**	.063
	3	.402***	.462***	.015	-.143**		.169**	.162*	.030	-.001	.349***	.321***	.092*	.015
Neurot.	1	.304***	-.304***	.043	-.002		.252***	-.239**	-.088	.031	.306***	-.308***	.021	-.033
	3	.503***	-.514***	.094*	-.050		.422***	-.406***	-.183***	-.060	.506***	-.491***	-.072	-.030
		Intelligence, intellectual style, and self-control traits												
Consc.	1	.458***	.451***	.089[a]	.069		.325***	.315***	.200***	.031	.467***	.423***	.150**	-.022
	3	.586***	.589***	-.025	-.075		.381***	.392***	.171***	.093	.577***	.547***	.133***	.093*
Open.	1	.220***	.200***	.083	.010		.140	.139[a]	.101[a]	.050	.238***	.183***	.127*	-.031
	3	.365***	.322***	.081	-.064		.155**	.218***	.057	.137*	.344***	.300***	.123**	.124**
Vocab.	2	.461***	-.098*	.464***	-.017		.389***	-.272***	.045	.122	.339***	-.099*	.323***	
	3	.497***	.535***	-.056	.077		.356***	-.316***	.142**	.041	.251***	-.050	.187***	.179***
		Workplace Behavior Traits												
OCB	2	.177**	.171***	.005	-.023		.154*	.245**	-.002	.174*	.181**	.123***	.037	.024
	3	.195***	.191***	-.101*	.022		.192***	.119	-.034	-.090	.224***	.205***	-.125**	-.034
CWB	2	.326***	-.221***	-.145**	.115*		.145*	-.163*	-.089	-.132	.321***	-.217***	-.216***	.014
	3	.337***	-.215***	-.116*	.158***		.061	.001	-.060	.003	.296***	-.139**	-.232***	.009
WDQ	2	.317***	.243***	.148**	-.056		.164*	.233**	.047	.210*	.289***	.256***	.136**	.052
	3	.335***	.314***	.038	-.055		.148*	.138*	.012	-.011	.301***	.297***	.039	.049

a. < .1
buffer text in between

The three models were kept separate, rather than combining all scores into one regression, owing to the issues of multicollinearity. If, for example, the original scores are used in conjunction with the scores for either DIZ or regression scores for confidence and accuracy, the scores—especially those for overconfidence and the SEPI, will exhibit multicollinearity. (Combining SEPI, TOPI, and either DIZs or regressions scores, the residualized accuracy scores were weighted most highly for all 10 regressions across Studies 1 and 3 related to the Big Five (i.e., 5 x 2), and additional variables exhibited multicollinearity issues).

Interpretation of Results. Examining Table 1, the strongest overall relations between the criterion scales and sets of variables arise for Models 1 and 3—that is, from employing the original SEPI and TOPI scores or from employing the residual scores for Overconfidence and Accuracy. The DIZ versions of the overconfidence and accuracy scores are substantially poorer overall predictors of this particular set of criteria. Although the DIZs are poorer as a group, the DIZ accuracy scores appear closely related to the variables of conscientiousness, vocabulary, and possibly also agreeableness, than the original TOPI scores alone. From a skeptical position, one might argue that is because they draw on some variance from the SEPI to do so; from a more optimistic perspective, one might say that they add prediction because, as absolute values of difference scores, they add substantially new information that neither the SEPI or TOPI can do alone.

Appendix C Table 2

Comparison Regressions on Criterion Variables Using Three Models, Each Employing Original SEPI and TOPI or Calculated Scores of Accuracy and Overconfidence, Standardized Betas

		Model 1 Original SEPI and TOPI Scores					Model 4 Original SEPI and TOPI Scores Plus DIZ Accuracy				Model 5 Original SEPI and TOPI Scores Plus Residual Accuracy			
Variable	Std.	Mult. R	SEPI[b]	TOPI[b]	Interact		Mult. R	SEPI	TOPI	Acc	Mult. R	SEPI	TOPI	Acc
		Socio-affective traits												
Extrav.	1	.389***	.389***	.039	.046		.386***	.383***	.029	.001	.390***	.391***	.069	-.065
	3	.317***	.327***	-.130**	.005		.317***	.334***	-.130**	-.020	.317***	.324***	-.140**	.017
Agree.	1	.354***	.346***	.080	.062		.352***	.334***	.048	.049	.362***	.324***	-.012	.127[a]
	3	.402***	.462***	.015	-.143**		.390***	.381***	.085[a]	-.094*	.380***	.352***	.075	.002
Neurot.	1	.304***	-.304***	.043	-.002		.305***	-.302***	.050	-.019	.305***	-.305***	.036	.012
	3	.503***	-.514***	.094*	-.050		.505***	-.499***	.127**	-.072[a]	.505***	-.505***	.151**	-.075
		Intelligence, intellectual style, and self-control traits												
Consc.	1	.458***	.451***	.089[a]	.069		.460***	.434***	.042	.085[a]	.466***	.427***	-.016	.144*
	3	.586***	.589***	-.025	-.075		.582***	.581***	.008	-.004	.584***	.568***	-.016	.054
Open.	1	.220***	.200***	.083	.010		.223***	.195***	.064	.045	.237***	.187***	.010	.114[a]
	3	.365***	.322***	.081	-.064		.370***	.342***	.118**	-.092*	.361***	.310***	.100*	.020
Vocab.	2	.461***	-.098*	.464***	-.017		.471***	-.085[a]	.498***	-.070	.467***	-.096*	.460***	.010
	3	.497***	.535***	-.056	.077		.492***	-.051	.502***	.013	.499***	-.026	.545***	-.098*
		Workplace Behavior Traits												
OCB	2	.177**	.171***	.005	-.023		.178**	.179***	.025	-.035	.177**	.172***	-.014	.035
	3	.195***	.191***	-.101*	.022		.198***	.207***	-.106*	-.043	.221***	.221***	-.056	-.127*
CWB	2	.326***	-.221***	-.145**	.115*		.307***	-.240***	-.183***	.027	.325***	-.230***	-.067	-.151*
	3	.337***	-.215***	-.116*	.158***		.307***	-211***	-.188***	.045	.329***	-.165***	-.120*	-.147**
WDQ	2	.317***	.243***	.148**	-.056		.317***	.260***	.187***	.036	.312***	.250***	.153*	.012
	3	.335***	.314***	.038	-.055		.342***	.337***	.071	-.094*	.334***	.320***	.085[a]	-.054

a. $p < .10$

One way to solve the issue is to add the accuracy term alone into the regression. . To allow for direct comparison, we have removed the interaction terms from Models 4 and 5, replacing them with the accuracy scores. When we do that, we obtain the following for Models 4 and 5 (Appendix C, Table 2). There is some conceptual and (potentially empirical) advantage to this approach, with the accuracy scores adding to the predictions in, arguably, the same or more number of cases as the interaction terms had. We further spot-checked three regressions in Study 2, adding the interaction terms back in; of the three analyses, the reintroduction of the interaction

term led to further improvements in one prediction. That is, when we added in the original Model 1 interaction term to Model 5, there is little difference for the behavior-at-work variables—except for an improvement of CWB for a multiple *R* from .325 to .342 in Study 2 (the only analysis for which we tested this), with all terms in the CWB regression significant.