International Journal of Communication 18(2024), 3922-3943 1932–8036/20240005

A Comparison of Four Approaches to Modeling Information Insufficiency

PENGYA AI♦
Nanyang Technological University, Singapore

SONNY ROSENTHAL[1]
Singapore Management University, Singapore

Information insufficiency, or the disparity between the level of knowledge needed to confidently judge an issue and the perceived level of current knowledge, is a key motivator of risk information seeking and processing. This study compared 4 approaches to modeling information insufficiency within the planned risk information seeking model. These approaches included the raw difference score, regression approach, partial variance score, and direct measure. Statistical modeling used data from large samples in Singapore (n = 2,124) and the United States (n = 2,125). The results of ordinary least squares regression analysis and structural equation modeling pointed to several issues. First, while the raw difference score is conceptually straightforward, it is susceptible to omitted variable bias when constructing explanatory models. The regression method is effective for data sets with low multicollinearity, while high multicollinearity warrants the analysis of partial variance. The direct measure, though simple, is prone to common method bias. Researchers should use the regression approach or partial variance score after assessing the degree of multicollinearity in their data sets.

Keywords: difference scores, differentials, information insufficiency, information seeking

Information insufficiency is a key variable in theories of information seeking, including the risk information seeking and processing (RISP) model (Griffin, Dunwoody, & Neuwirth, 1999) and the planned risk information seeking model (PRISM; Kahlor, 2010). Many studies suggest it is essential in predicting information-seeking behavior (Hwang & Jeong, 2016). Those studies present consistent conceptualizations and theorizations about information insufficiency, but there are different operationalizations and analytical approaches that can produce incompatible results and hamper meta-analyses.

Pengya Ai: pengya001@e.ntu.edu.sg
Sonny Rosenthal: srosenthal@smu.edu.sg
Date submitted: 2024-02-20

[1] This research received funding support from an AcRF Tier 1 grant by the Singapore Ministry of Education (RG121/20).

This is not the first research effort to note this problem. Rosenthal (2013) compared three approaches to statistically model a common two-part measure of information insufficiency. Those approaches focus on the difference between an information sufficiency threshold and perceived current knowledge. Based on a simulation study, he recommended using the analysis of partial variance, which involves partialing out the variance in the sufficiency threshold that is related to current knowledge and using the residual variance as an indicator of information insufficiency. He argued that approach is most appropriate when current knowledge is highly correlated with other predictors of information seeking, for example, informational subjective norms.

The current study addresses three lingering research gaps. First, the data simulation in Rosenthal (2013) is not representative of real-world data, mainly because the variables he created had exactly normal distributions. Second, the data simulation did not situate information insufficiency within a theoretical framework, and it remains unknown how the different approaches might affect theoretical conclusions drawn from real-world data sets. Third, some studies have used a direct measure of information insufficiency in which respondents indicate to what extent they feel their knowledge is sufficient for their purposes (Rosenthal, 2011; Trumbo, 2002). There has not been research comparing that direct measure with the two-part measure. The current study seeks to fill these gaps by analyzing data from a cross-sectional survey in Singapore and the United States. By using real data from two countries, this study can provide more generalizable findings and practical recommendations for future analyses.

Information Insufficiency

Information insufficiency is a key motivator of risk information seeking and processing (Hwang & Jeong, 2016; Kahlor, 2010). It is the disparity between the level of knowledge needed to pass confident judgment about an issue—also called the *sufficiency threshold*—and the perceived level of current knowledge (Griffin et al., 1999; Griffin, Neuwirth, Dunwoody, & Giese, 2004). Theories of risk information seeking drew that concept from the sufficiency principle of the heuristic-systematic model of persuasion (Chaiken, 1980; Chaiken & Trope, 1999). According to that model, people are cognitive misers, investing mental effort to process information only when necessary (Eagly & Chaiken, 1993). The sufficiency principle states that individuals will invest that effort to attain what they feel is a sufficient level of confidence in accomplishing their processing goals (Chen & Chaiken, 1999; Eagly & Chaiken, 1993). This principle can explain information seeking as a function of information insufficiency. When individuals feel their knowledge falls short of a sufficiency threshold, they experience information insufficiency, which can motivate them to seek more information (Griffin et al., 2004; Hwang & Jeong, 2016).

Information insufficiency is related to several factors, including attitude toward seeking, seeking-related subjective norms, perceived seeking control, and negative affect (Kahlor, 2010). Attitude toward seeking refers to favorable or unfavorable evaluations of the information-seeking behavior. Seeking-related subjective norms are the perceived behaviors or expectations of others that create social pressure to seek information. Perceived seeking control is the perception individuals have of their abilities and efficacy to seek information. Finally, negative affect is an emotional response, such as worry and anxiety, to a perceived risk. According to the PRISM, three of these variables—attitude toward seeking, seeking-related subjective norms, and negative affect—are positively related to information insufficiency and seeking intention (Kahlor, 2010; Kahlor et al., 2020). These linkages are intuitive, as believing that information seeking is useful (an attitude) and prevalent (a subjective norm) may increase the sufficiency threshold. Believing that a personal or environmental risk is worrisome can also increase the sufficiency threshold specific to that risk. In contrast, perceived seeking control positively predicts information-seeking intention but negatively predicts information insufficiency (Kahlor, 2010). This is also reasonably intuitive, as being capable of seeking information can motivate information seeking, but individuals who feel they have control over their information seeking may tend not to feel their knowledge is lacking. Figure 1 summarizes these theoretical linkages.

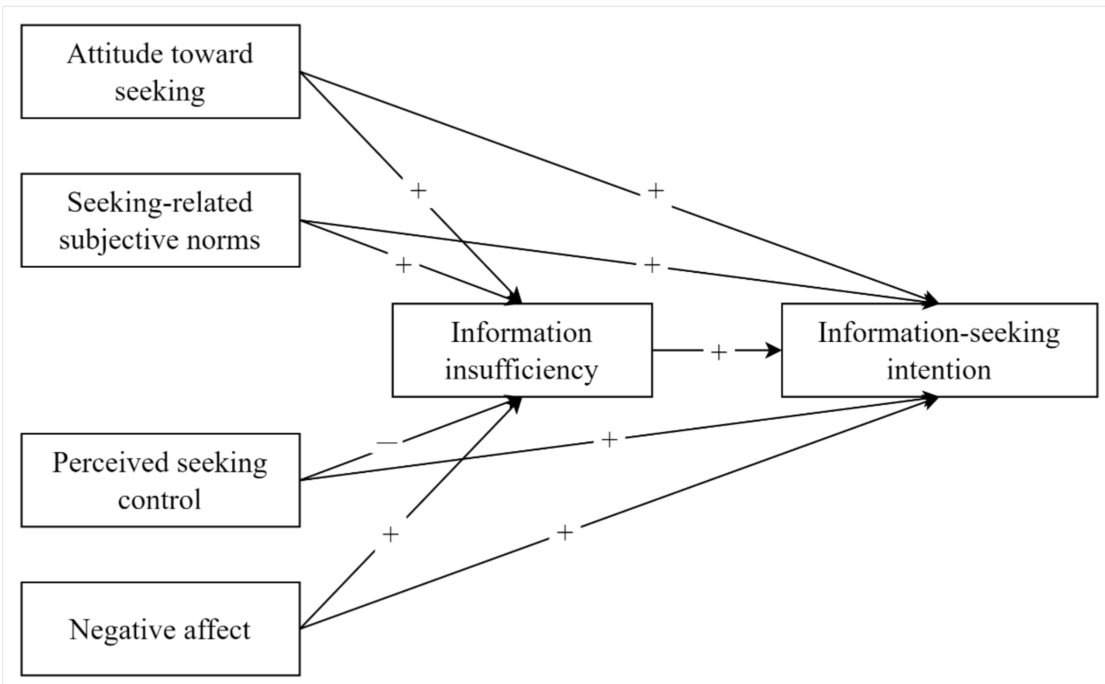

Figure 1. The PRISM (truncated).

The Operationalization of Information Insufficiency

Two-Part Measure

Most studies of information insufficiency have used a two-part measure of the information sufficiency threshold and current knowledge, which Griffin et al. (2004) developed. The measure of information sufficiency threshold asks respondents to report their desired knowledge or the information they think they need to achieve a comfortable understanding or pass confident judgment about a topic. The measure of perceived current knowledge asks respondents to rate their current knowledge of the topic. Usually, both measures are on a scale of 0 (*knowing nothing*) to 100 (*knowing everything*). Scholars have been consistent in the use of these measures, with only slight variations in the question wording; however, they have employed different approaches in the statistical modeling and analysis of information insufficiency. At least three approaches appear in the literature, including raw difference scores (e.g., Ahn & Kahlor, 2020), a regression approach (e.g., Hwang & Jeong, 2016; Kahlor et al., 2020; Yang & Kahlor, 2012), and the analysis of partial variance (e.g., Ahn & Kahlor, 2022; Ho, Detenber, Rosenthal, & Lee, 2014; Kim & Lai, 2019).

Raw Difference Score

Computing a raw difference score is perhaps the most intuitive way to calculate information insufficiency. This method obtains information insufficiency by subtracting the current knowledge score from the sufficiency threshold score. For example, individuals who report a current knowledge of 50 and a sufficiency threshold of 80 have an information insufficiency of 30.

Although a raw difference score is conceptually easy to understand, there are several pitfalls in using it to represent information insufficiency. Griffin et al. (2004) argued that a raw difference score can multiply reliability problems and is prone to floor and ceiling effects. Cohen, Cohen, West, and Aiken (2003) noted that a raw difference score implies that the unstandardized regression slope relating the two variables is exactly 1, which is unrealistic in most instances of social scientific research and can bias the estimated difference. Further, Rosenthal (2013) showed that the raw difference score could bias explained variance when it serves as either an independent or dependent variable.

Regression Approach

Noting these deficiencies of using raw difference scores, scholars have developed two alternative approaches to estimate information insufficiency. One approach uses current knowledge as a control variable in regression models. Early descriptions of this approach conveyed it in two steps. The first step regresses sufficiency threshold on current knowledge. The unexplained variance, which is the variance of the sufficiency threshold unrelated to current knowledge, is a reasonable indicator of information insufficiency (Griffin et al., 2004). The second step adds other predictors of sufficiency threshold. Since those variables predict the unexplained variance from the first step, there is a good argument that the slopes show the effects of those variables on information insufficiency (Kahlor et al., 2020; Lu, Group, Winneg, Jamieson, & Albarracin, 2020). This is the procedure when information insufficiency is the dependent variable. When it is an independent variable, the first step regresses the dependent variable (e.g., seeking intention) on current knowledge, and the second step adds sufficiency threshold to predict the residual variance in the dependent variable. In both models, separating the steps is unnecessary, but it helps to explain how the approach works.

Although this method can avoid the floor and ceiling effects, low reliability, and bias that using raw difference scores creates, the regression approach could cause problems if there is multicollinearity between current knowledge and other predictor variables (Rosenthal, 2013). This is because the second step does not account for the effects of other predictors on the unexplained variance of current knowledge. When the additional predictors are included in the model, the slope associated with current knowledge changes as a function of its relationship with those predictors. Thus, if there is multicollinearity, it creates a problem for interpreting effects on and of information insufficiency.

Partial Variance

The second approach to address the deficiencies of raw difference scores involves computing the partial variance of the sufficiency threshold. This approach requires two steps. The first step is the same as the regression approach, regressing sufficiency threshold (*ST*) on current knowledge (*CK*). The second step computes information insufficiency (*II*) as partial variance, using the regression slope (*B*) from the first step to weight current knowledge, which is subtracted from the sufficiency threshold. Equation 1 shows the computation.

$$II = TH - B \cdot CK \tag{1}$$

This approach avoids the multicollinearity problem of the regression approach while maintaining all its benefits relative to the use of raw difference scores (Rosenthal, 2013). It is similar to the regression approach because it analyzes the variance of the sufficiency threshold controlling for current knowledge, which is the key benefit of both approaches. However, it avoids the multicollinearity problem of the regression approach because the second step removes the variance related to current knowledge. Thus, the relationship between information insufficiency and other predictors will be independent of the relationship between current knowledge and those other predictors.

Direct Measure

The two-part measure of information insufficiency parallels the two-part definition of the concept in terms of the sufficiency threshold and current knowledge. However, it is also possible to measure information insufficiency directly, which can simplify its analysis. Rosenthal (2011) developed a three-item direct measure of information insufficiency based on a prior single-item measure (Trumbo, 2002). The items ask survey respondents if the information they currently have meets all of their needs for knowing about a topic (reverse-coded), their current level of knowledge about that topic is sufficient (reverse-coded), and if they have not received enough information about that topic.

The current study draws comparisons among these four approaches to modeling information insufficiency as an independent variable predicting information-seeking intention and, separately, a dependent variable predicted by several covariates. Consistent with the PRISM (Kahlor, 2010), the covariates include attitude toward seeking, seeking-related subjective norms, perceived seeking control, and negative affect. Figure 2 shows conceptual diagrams of the four approaches.

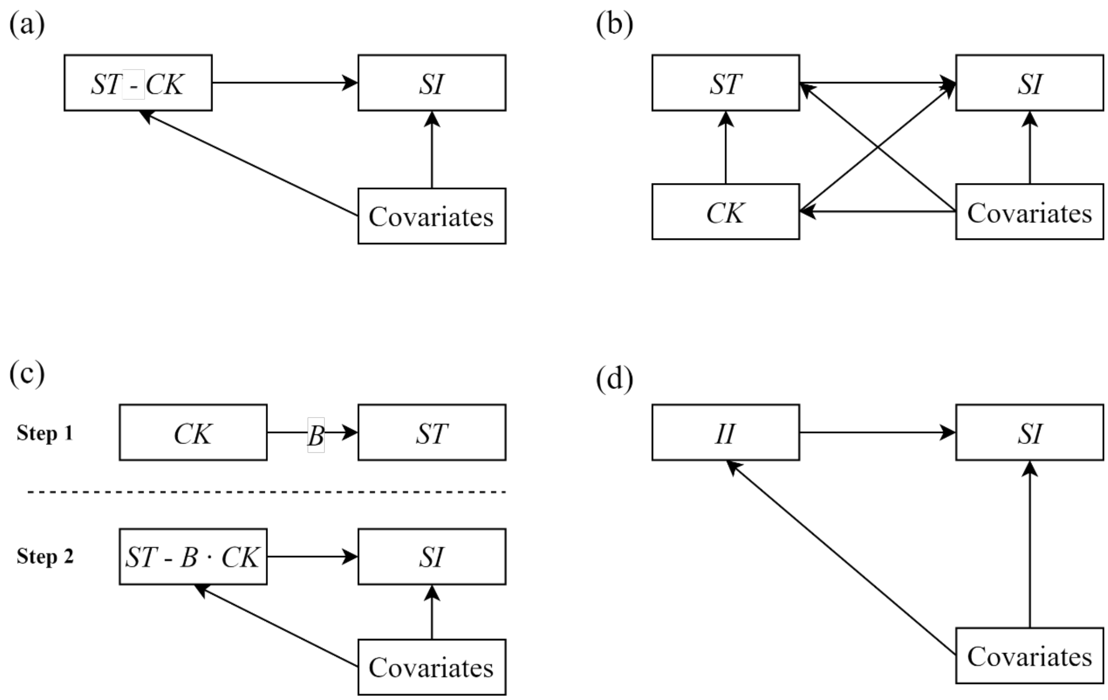

Figure 2. Four models of information insufficiency.

Note. (a) raw difference score, (b) regression approach, (c) partial variance approach, and (d) direct measure. *SI* = seeking intention. *ST* = sufficiency threshold. *CK* = current knowledge. *II* = information insufficiency.

Method

Study Context

We modeled information insufficiency about climate change. Climate change is a complex issue that poses serious threats to humanity (Hunt & Watkiss, 2011; Wheeler & von Braun, 2013). Adequately responding to it will require accurate and timely information about its causes and effects (Kahlor et al., 2020). Information seeking can improve public knowledge and support effective policymaking (Li, Qin, Quan, & Tan-Soo, 2023). It is also the context of many studies on risk information seeking (e.g., Ho et al., 2014; Kahlor et al., 2020; Yang, Kahlor, & Li, 2013), making it a fitting topic of this methodological work.

We conducted this study in Singapore and the United States. These countries have distinct ideologies, politics, and public opinions on the environment. For example, unlike individuals in the United States, Singaporeans received relatively consistent information on climate change (Detenber, Rosenthal, Liao, & Ho, 2016). The Singapore government has passed laws regulating misinformation, which may include misinformation about climate change (Singapore Legal Advice [SLA], 2022). Similarly restrictive laws do not exist in the United States. Also, more than 90% of Singaporeans believe climate change is happening and is because of human activities, which is markedly higher than in the United States (Center for Climate Change Communication [CCCC], 2022; Mandloi, 2019). These differences may lead to diverging perceptions of climate change and information-seeking behaviors. Therefore, the cross-national comparison can increase the generalizability of this study.

Data Source

The Institutional Review Board at Nanyang Technological University approved this study, which included documented informed consent. We sampled survey respondents in Singapore (n = 2,124) and the United States (n = 2,125) over three weeks in April and May 2022 using online research panels of Rakuten Insights. Panelists were randomly sampled from the panels, but with quotas for age and sex corresponding to population demographics. In both countries, Rakuten Insights calculated a response rate of 20% following the RR2 response rate definition by the American Association for Public Opinion Research (AAPOR, 2023).

Women comprised 52% of the Singapore sample and 53% of the United States sample. The Singapore sample had a median age of 39 years (M = 40.86, SD = 13.26), compared with 41 years in the United States sample (M = 43.75, SD = 14.99). The median educational attainment was a bachelor's degree in Singapore and an associate's degree in the United States. The median income bracket was SGD $7,000 to $8,999 in Singapore and USD $5,000 to $5,999 in the United States. These figures generally match with census figures, except that the median education levels in both samples were higher than the census (Census Bureau, 2020, 2021a, 2021b; Singapore Department of Statistics [SDS], 2021a, 2021b).

Measurements

We adapted measures from Kahlor (2010), Kahlor et al. (2020), and Rosenthal (2011). Table 1 contains the means and standard deviations of composite variables.

Current Knowledge and Sufficiency Threshold

Respondents rated their current knowledge about the causes, effects, and science of climate change on a scale from 0 (*knowing nothing*) to 100 (*knowing everything they could possibly know*). The items had acceptable reliability (Cronbach's α = 0.91). Using the same 0 to 100 scale, respondents reported how much information they felt they would need to have an informed opinion about the causes, effects, and science of climate change. The items had acceptable reliability (Cronbach's α = 0.90).

Raw Difference and Partial Variance

To obtain raw difference scores, we subtracted current knowledge from the sufficiency threshold for the causes, effects, and science of climate change. The three raw difference scores had acceptable reliability (Cronbach's α = 0.85). To obtain partial variance scores, we regressed the sufficiency threshold on current knowledge separately in the Singapore and United States samples. Notably, the slopes were not different between countries with respect to the causes (p = 0.61), effects (p = 0.63), and science (p = 0.93) of climate change. Then we subtracted the regression-weighted current knowledge from the sufficiency threshold for the causes, effects, and science of climate change. The three partial variance scores had acceptable reliability (Cronbach's ɑ = 0.87).

Direct Measure

Respondents indicated their level of agreement that "The information I have at this time meets all of my needs for knowing about climate change"; "My current level of knowledge about climate change is sufficient"; and "I have not received enough information about climate change." Response options ranged from 1 (*strongly disagree*) to 5 (*strongly agree*). We removed the third item because of low reliability (Cronbach's ɑ = 0.46). The remaining two items, which were reverse-coded, had acceptable reliability (Spearman-Brown coefficient = 0.78).

Information-Seeking Intention

Respondents indicated how much information they intended to seek about the causes, effects, and science of climate change in the next month. Response options ranged from 1 (*none at all*) to 5 (*everything possible*), and the items had acceptable reliability (Cronbach's ɑ = 0.93).

Attitude Toward Seeking

Respondents indicated their instrumental attitude toward seeking on four semantic differential items comprising opposing pairs of adjectives. The pairs of adjectives included *worthless*: *valuable*, *harmful*: *beneficial*, *unhelpful*: *helpful*, and *foolish*: *wise*. Responses were scored on a 5-point scale, where a higher score indicated a more positive attitude, and the items had acceptable reliability (Cronbach's ɑ = 0.89).

Seeking-Related Subjective Norms

Respondents indicated their level of agreement that "people I spend most of my time with are likely to seek information related to climate change"; "people who are close to me expect me to seek information about climate change"; and "most people who are important to me think that I should seek information about climate change." The first item measured a descriptive norm pertaining to the behavior of others, while the latter two items measured an injunctive norm pertaining to social expectations. Response options ranged from 1 (*strongly disagree*) to 5 (*strongly agree*), and the items had acceptable reliability (Cronbach's ɑ = 0.86).

Perceived Seeking Control

Respondents indicated their level of agreement that "it is easy to find information about climate change"; "I know how to search for information about climate change"; "I know where to look for information about climate change;" and "I can readily access all the information I need about climate change." Response options ranged from 1 (*strongly disagree*) to 5 (*strongly agree*), and the items had acceptable reliability (Cronbach's ɑ = 0.85).

Negative Affect

Respondents indicated their level of agreement that "I feel worried about climate change"; "I feel anxious about climate change"; and "I feel concerned about climate change." Response options ranged from 1 (*strongly disagree*) to 5 (*strongly agree*), and the items had acceptable reliability (Cronbach's α = .90).

Table 1. Descriptive Statistics of All Variables.

Variables	***M***	***SD***	**Min**	**Max**	**Skewness**	**Kurtosis**
Information-seeking intention	3.11	1.08	1	5	−0.23	2.41
Sufficiency threshold	64.85	24.42	0	100	−0.57	2.74
Current knowledge	56.85	24.31	0	100	−0.25	2.41
Partial variance	38.85	21.09	−45.60	100	−0.32	3.88
Raw difference score	8.00	24.22	−100	100	0.27	5.80
Direct measure	3.25	0.95	1	5	−0.13	2.52
Perceived seeking control	3.75	0.74	1	5	−0.49	3.55
Seeking-related subjective norms	3.14	0.96	1	5	−0.22	2.64
Attitude toward seeking	4.03	0.98	1	5	−0.86	3.11
Negative affect	3.79	0.99	1	5	−0.94	3.61

Analytical Approach

We used two different approaches to estimate the four models of information insufficiency. First, we conducted ordinary least squares (OLS) regression, which is common in the social sciences (e.g., Ahn & Kahlor, 2020, 2022; Kim & Lai, 2019). Second, we conducted structural equation modeling (SEM) with maximum likelihood estimation. This latter method analyzes latent variables, which reflect the common variances among variable indicators (Rosenthal, 2017), and is increasingly common in communication research, particularly in survey research (e.g., Ho et al., 2014; Kahlor et al., 2020; Lu et al., 2020). We used both methods because the choice of estimator in a regression model may influence the results (Hayes, Montoya, & Rockwood, 2021) and we wanted our findings to have wide applicability. We used Stata version 17 to estimate the OLS regression models and Mplus version 8.3 to estimate the SEM models.

Our SEM models established measurement and structural models. The measurement model constructed the latent variables and the structural model specified paths among those variables. We assessed model fit using the cutoff criteria of Hu and Bentler (1999), who recommended that CFI should be higher than 0.95, RSMEA should be lower than 0.06, and SRMR should be lower than 0.08. Since our data were collected in two countries, we estimated multigroup measurement models. That analysis involved establishing measurement invariance at three levels, indicating increasing consistency of measurement between the two countries. The first level, configural invariance, establishes that the same measurement items are related to the same latent variables. The second level, metric invariance, constrains the factor loadings to be equal between countries, allowing between-country comparisons of regression slopes. The third level, scalar invariance, additionally constrains the item intercepts to be equal, allowing between-country comparisons of factor mean scores in addition to regression slopes. Following the criteria of Chen (2007), we found support for scalar measurement invariance in all four models (see Table 2).

We tested the models of information insufficiency using a two-step hierarchical regression analysis. In the first step, we examined the relationship between information insufficiency and seeking intention without the covariates. This is because potential multicollinearity between information insufficiency and other covariates may change the relationship between information insufficiency and information seeking (Rosenthal, 2013). In the second step, we included the covariates to fully model the PRISM and provide a more comprehensive view of the relationship between information insufficiency and information seeking.

Table 2. Measurement Invariance.

Model fit	**χ^2**	***df***	**CFI**	**RSMEA [90% CI]**	**SRMR**	**ΔCFI**	**ΔRSMEA**	**ΔSRMR**
The measurement model with the raw difference score								
Configural model	1190.08	310	0.98	0.04 [0.03, 0.04]	0.03			
Metric model	1203.14	324	0.98	0.04 [0.03, 0.04]	0.03	< 0.001	−0.001	< 0.001
Scalar model	1243.55	338	0.98	0.04 [0.03, 0.04]	0.03	−0.001	< 0.001	0.001
The measurement model with the regression approach								
Configural model	2199.88	418	0.97	0.05 [0.04, 0.05]	0.03			
Metric model	2220.50	434	0.97	0.04 [0.04, 0.05]	0.03	< 0.001	−0.001	0.001
Scalar model	2299.40	450	0.97	0.04 [0.04, 0.05]	0.03	−0.001	< 0.001	< 0.001
The measurement model with the partial variance								
Configural model	1097.80	310	0.99	0.04 [0.03, 0.04]	0.03			
Metric model	1118.18	324	0.99	0.03 [0.03, 0.04]	0.03	−0.001	−0.001	0.001
Scalar model	1163.06	338	0.99	0.03 [0.03, 0.04]	0.03	< 0.001	< 0.001	< 0.001
The measurement model with the direct measure								
Configural model	933.77	274	0.99	0.03 [0.03, 0.04]	0.03			
Metric model	945.78	287	0.99	0.03 [0.03, 0.04]	0.03	< 0.001	−0.001	0.001
Scalar model	1050.80	300	0.99	0.03 [0.03, 0.04]	0.03	−0.002	0.001	0.001

Note. Each multigroup measurement model includes one of the four measures of information insufficiency, information-seeking intention, and the four covariates. The rightmost three columns show the changes in model fit going from a less constrained model to a more constrained model. χ^2 = Chi-square. *df* = degrees of freedom. CFI = comparative fit index. RMSEA = root mean square error of approximation. SRMR = standardized root mean square residual. ΔCFI = change in CFI. ΔRSMEA = change in RMSEA. ΔSRMR = change in SRMR.

Results

Models Sans Covariates

We first examined the models without covariates. The four SEM models had acceptable fit (see Table 3).

Table 3. Fit of SEM Sans Covariates.

Model	χ2 (*df*)	CFI	RSMEA [90% CI]	SRMR
a. Raw difference score	80.64 (24)	1.00	0.03 [0.03, 0.04]	0.02
b. Regression approach	853.97 (60)	0.98	0.08 [0.07, 0.08]	0.04
c. Partial variance	127.79 (24)	0.99	0.05 [0.04, 0.05]	0.02
d. Direct measure	104.45 (14)	0.99	0.06 [0.05, 0.07]	0.02

Note. χ^2 = Chi-square. *df* = degree of freedom. CFI = comparative fit index. RSMEA = root mean square error of approximation. CI = confidence interval. SRMR = standardized root mean square residual.

Table 4 contains the standardized coefficients from the OLS regression and SEM models. Contradicting the PRISM, the raw difference score was negatively related to information-seeking intention. In contrast, the other three methods resulted in a positive relationship between information insufficiency and seeking intention. Using the regression approach or partial variance produced small effects. Using the direct measure of information insufficiency resulted in a medium effect, and it appears that effect was larger in the SEM model than in the OLS model.

Table 4. The Relationship Between Information Insufficiency and Information-Seeking Sans Covariates.

		Singapore		United States	
Model	Paths	OLS	SEM	OLS	SEM
a. Raw difference score	*II* → *SI*	-0.10^{***}	-0.11^{***}	-0.10^{***}	-0.09^{***}
	R^2	0.01	0.01	0.01	0.01
b. Regression approach	*CK* → *SI*	0.35^{***}	0.35^{***}	0.38^{***}	0.41^{***}
	ST → *SI*	0.12^{***}	0.11^{***}	0.20^{***}	0.19^{***}
	R^2	0.18	0.17	0.27	0.29
c. Partial variance	*II* → *SI*	0.12^{***}	0.14^{***}	0.20^{***}	0.22^{***}
	R^2	0.02	0.02	0.04	0.05
d. Direct measure	*II* → *SI*	0.38^{***}	0.45^{***}	0.33^{***}	0.39^{***}
	R^2	0.14	0.20	0.11	0.15

Note. Estimates are standardized regression slopes. *II* = information insufficiency. *SI* = seeking intention. *ST* = sufficiency threshold. *CK* = current knowledge. R^2 = explained variance of seeking intention. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. The relationship between *CK* and *ST* were omitted in the table for clarity.

Models With Covariates

Next, we added covariates as predictors of both information insufficiency and information-seeking intention. Table 5 summarizes the fit of the four SEM models, which all had acceptable fit.

Table 5. Model Fits of Models With Covariates.

Method	**χ2 (*df*)**	**CFI**	**RSMEA [90% CI]**	**SRMR**
a. Raw difference score	1243.55 (338)	0.98	0.04 [0.03, 0.04]	0.03
b. Regression approach	2299.40 (450)	0.97	0.04 [0.04, 0.05]	0.03
c. Partial variance	1163.06 (338)	0.99	0.03 [0.03, 0.04]	0.03
d. Direct measure	1050.80 (300)	0.99	0.03 [0.03, 0.04]	0.03

Note. χ^2 = Chi-square. *df* = degree of freedom, CFI = comparative fit index. RSMEA = root mean square error of approximation. CI = confidence interval. SRMR = standardized root mean square residual.

Raw Difference Score

Using the raw difference score (Table 6), information insufficiency was positively predicted by attitude toward seeking and negatively predicted by perceived seeking control. Those paths were consistent with the PRISM. Inconsistent with the PRISM, it was negatively predicted by subjective norms and unrelated to negative affect. Seeking intention was positively predicted by subjective norms and negative affect. Except for the SEM model in the United States, it was also positively predicted by perceived seeking control. In contrast, it was unrelated to information insufficiency and attitude toward seeking. The pattern of results was consistent between Singapore and the United States.

Table 6. Standardized Coefficients Using the Raw Difference Score.

	Singapore		**United States**	
Paths	**OLS**	**SEM**	**OLS**	**SEM**
ATT → *II*	0.22^{***}	0.24^{***}	0.23^{***}	0.26^{***}
SN → *II*	-0.17^{***}	-0.16^{***}	-0.09^{***}	-0.07^{*}
PC → *II*	-0.18^{***}	-0.24^{***}	-0.24^{***}	-0.30^{***}
NA → *II*	0.03	0.05	0.04	0.03
R^2	0.09	0.11	0.10	0.13
II → *SI*	−0.002	0.02	−0.01	0.00
ATT → *SI*	−0.01	−0.02	−0.001	−0.01
SN → *SI*	0.53^{***}	0.59^{***}	0.47^{***}	0.54^{***}
PC → *SI*	0.07^{***}	0.06^{**}	0.07^{***}	0.04
NA → *SI*	0.20^{***}	0.20^{***}	0.29^{***}	0.28^{***}
R^2	0.43	0.52	0.49	0.58

Note. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. *ATT* = attitude. *SN* = seeking-related subjective norms. *PC* = perceived seeking control. *NA* = negative affect. *II* = information insufficiency. *SI* = seeking intention. R^2 = explained variance.

Regression Approach

Using the regression approach (Table 7), information insufficiency was positively predicted by attitude toward seeking and negative affect, and negatively predicted by perceived seeking control. Whereas in the United States, it was also positively predicted by subjective norms, that relationship was negative in Singapore, contradicting the PRISM. Seeking intention was positively predicted by information insufficiency, subjective norms, and negative affect. Perceived seeking control was also a positive predictor, but only in the OLS model. Inconsistent with the PRISM, attitude toward seeking was a negative predictor in Singapore, and that relationship was nonsignificant in the United States.

Table 7. Standardized Coefficients Using the Regression Approach.

	Singapore		United States	
Paths	**OLS**	**SEM**	**OLS**	**SEM**
CK → *ST*	0.43***	0.46***	0.46***	0.51***
ATT → *ST*	0.25***	0.26***	0.22***	0.25***
SN → *ST*	−0.05**	−0.07**	0.07**	0.08**
PC → *ST*	−0.06**	−0.09***	−0.10***	−0.15***
NA → *ST*	0.12***	0.14***	0.10***	0.08*
R^2	0.33	0.37	0.35	0.39
CK → *SI*	0.12***	0.09***	0.14***	0.12***
ST → *SI*	0.11***	0.12***	0.09***	0.09***
ATT → *SI*	−0.04*	−0.05*	−0.02	−0.03
SN → *SI*	0.50***	0.58***	0.41***	0.49***
PC → *SI*	0.04*	0.03	0.04*	0.01
NA → *SI*	0.16***	0.16***	0.26***	0.26***
R^2	0.46	0.54	0.52	0.60

Note. *$p < .05$. **$p < .01$. ***$p < .001$. *ATT* = attitude. *SN* = seeking-related subjective norms. *PC* = perceived seeking control. *NA* = negative affect. *CK* = current knowledge. *ST* = information sufficiency threshold. *SI* = seeking intention. The relationships between *CK* and covariates were omitted in the table for clarity.

Partial Variance

Using the partial variance approach (Table 8), the prediction of information insufficiency followed the same pattern as the regression approach. It was positively predicted by attitude toward seeking and negative affect and negatively predicted by perceived seeking control. Subjective norms were a positive predictor in the United States and a negative predictor in Singapore. Seeking intention was positively predicted by information insufficiency, subjective norms, and perceived seeking control, and was unrelated to attitude toward seeking. Different from the regression approach, perceived seeking control was a significant predictor in both the OLS and SEM models.

Table 8. Standardized Coefficients Using Partial Variance.

	Singapore		United States	
Paths	**OLS**	**SEM**	**OLS**	**SEM**
ATT → *II*	0.28***	0.30***	0.26***	0.29***
SN → *II*	−0.06**	−0.07**	0.09***	0.12**
PC → *II*	−0.07**	−0.10***	−0.12***	−0.16***

NA → II	0.13^{***}	0.16^{***}	0.11^{***}	0.09^{*}
R^2	0.11	0.14	0.11	0.13
II → SI	0.09^{***}	0.10^{***}	0.08^{***}	0.08^{***}
ATT → SI	−0.04	−0.04	−0.02	−0.03
SN → SI	0.53^{***}	0.60^{***}	0.47^{***}	0.53^{***}
PC → SI	0.08^{***}	0.07^{**}	0.08^{***}	0.06^{**}
NA → SI	0.19^{***}	0.18^{***}	0.28^{***}	0.27^{***}
R^2	0.44	0.52	0.50	0.58

Note. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. *ATT* = attitude. *SN* = seeking-related subjective norms. *PC* = perceived seeking control. *NA* = negative affect. *II* = information insufficiency. *SI* = seeking intention.

Direct Measure

Using the direct measure (Table 9), information insufficiency was positively predicted by subjective norms. Contradicting the PRISM, it was negatively predicted by attitude toward seeking and positively predicted by perceived seeking control. Except for the OLS model in Singapore, it was also negatively predicted by negative affect. Seeking intention was positively related to subjective norms, perceived seeking control, and negative affect. It was unrelated to attitude toward seeking and was positively related to information insufficiency only in the OLS model in Singapore.

Table 9. Standardized Coefficients Using the Direct Measure.

	Singapore		**United States**	
Paths	**OLS**	**SEM**	**OLS**	**SEM**
ATT → *II*	-0.13^{***}	-0.17^{***}	-0.24^{***}	-0.28^{***}
SN → II	0.42^{***}	0.51^{***}	0.39^{***}	0.47^{***}
PC → II	0.37^{***}	0.48^{***}	0.44^{***}	0.55^{***}
NA → II	−0.02	-0.07^{*}	-0.09^{***}	-0.15^{***}
R^2	0.37	0.56	0.44	0.64
II → SI	0.05^{**}	−0.01	0.03	−0.04
ATT → SI	−0.00	−0.01	0.00	−0.02
SN → SI	0.50^{***}	0.60^{***}	0.46^{***}	0.56^{**}
PC → SI	0.05^{**}	0.06^{*}	0.06^{**}	0.07^{*}
NA → SI	0.20^{***}	0.20^{***}	0.29^{***}	0.27^{***}
R^2	0.44	0.52	0.49	0.58

Note. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. *ATT* = attitude. *SN* = seeking-related subjective norms. *PC* = perceived seeking control. *NA* = negative affect. *II* = information insufficiency. *SI* = seeking intention.

Discussion

This study used data from a large cross-national survey to compare four approaches to modeling information insufficiency in the PRISM. This extended previous work that used simulated data of the two-part measurement of information insufficiency. Current findings showed that the choice of measurement and estimation method affect statistical outcomes and theoretical implications. We discuss these findings and make recommendations with respect to theoretical consistency and statistical factors.

First, we discuss the three applications of the two-part measurement, which replicated Rosenthal's (2013) study by utilizing real data within a theoretical framework. Like Rosenthal (2013), we found an unusual relationship between the raw difference score and the intention to seek information. Our analysis showed a negative effect in the simple regression model and a nonsignificant effect in the multiple regression model. This seems because of the negative correlation between the raw difference score and current knowledge ($r_{\text{pooled}} = -0.49$, $p < .001$), which may have introduced omitted variable bias. This bias occurs when an important variable is left out of a predictive model, leading to distorted or biased estimates of the relationships between included variables (Wooldridge, 2015). In this case, the omitted variable was current knowledge. In contrast, the regression approach and analysis of partial variance accounted for the relationship between information insufficiency and current knowledge; thus, the same bias was less likely to arise. This observation can explain why, compared with the regression approach and partial variance, the raw difference score failed to reproduce theoretically consistent paths. It may be possible to use a raw difference score in a predictive model, but the model needs to control for current knowledge. Although such methodological amendment is beyond the scope of this study, we provide a detailed mathematical and empirical demonstration of it in the appendix.

Although the regression approach and analysis of partial variance produced results that were mostly consistent with the theoretical expectations, the former was subject to a special bias because of multicollinearity. This bias occurs when the model includes additional covariates that correlate with current knowledge. As a result, the regression of sufficiency threshold on current knowledge changes from step 1 to step 2 of the regression approach. As Rosenthal (2013) demonstrated, this does not happen with the analysis of partial variance. This difference explains why the two approaches yielded slightly divergent findings that also differed between the OLS regression and SEM models. For example, the relationship between perceived control and seeking intention was positive except when using SEM to conduct the regression approach. This highlights the importance of not only model specification but also the choice of model estimation. Though, given the generally small differences we observed, the current analysis does not clearly favor one specification and estimator over the others. The differences between the regression approach and analysis of partial variance were likely because of current knowledge having small to moderate correlations with the covariates. Researchers can expect larger disparities when the relationships between current knowledge and covariates approach unity, warranting the analysis of partial variance (Rosenthal, 2013). Other research has also noted the robustness of the analysis of partial variance to multicollinearity (Dougherty, 2011). As a general recommendation, researchers should examine the correlations between current knowledge and covariates before choosing a modeling approach.

Second, we discuss the use of the direct measure and how it compares to the analyses of the two-part measurement. Previous studies suggested the direct measure is a more convenient way to measure information insufficiency because it omits the need to model two separate measures and all the decisions such modeling entails (Rosenthal, 2011). However, until the current study, researchers had not statistically evaluated the direct measure in comparison to other modeling approaches. In our simple regression model, where information insufficiency predicted seeking intention, analysis of the direct measure resulted in more positive coefficients than analysis of any of the two-part measures. That observation lends support to the use of the direct measure.

However, after we added covariates, that model performed much worse than the regression approach or analysis of partial variance. A likely culprit behind this discrepancy is common method bias. The direct measure of information insufficiency used the same 5-point Likert scale as information-seeking intention, perceived seeking control, subjective norms, and negative affect. This is not an inherent feature of the direct measure, as researchers could simply use a different scale of measurement, for example, by asking about degrees of sufficiency on a scale from minus 10 to positive 10. Therefore, we cannot argue conclusively against using the direct measure. Instead, we call for further research to examine its explanatory power conditioned on the similarity of its measurement to the measurement of other variables.

Although we did not aim to validate the PRISM, we wish to emphasize one between-country difference that may have theoretical import. The relationship between subjective norms and information insufficiency was negative in Singapore and positive in the United States in the model using regression approach and partial variance. This difference could be because of cultural factors. For example, Singapore is a more collectivist culture, whereas the United States is a more individualist culture. Research has linked these cultural orientations with the influence of social norms (Huang, Leung, Eom, & Tam, 2022; Lapinski & Rimal, 2005). However, since we did not measure these or other cultural factors, further discussion of this would be speculative. Such an explanation would be a good focus of future research.

To conclude this work, we briefly summarize our thoughts about the different approaches to modeling information insufficiency and key takeaways from our findings. Whereas the raw difference score seems a straightforward representation of the construct, it is prone to omitted variable bias, at least in the way researchers commonly use it. The regression approach is optimal when there is low multicollinearity between current knowledge and other covariates. When there is high multicollinearity, the analysis of partial variance will produce the least biased estimates, but we are unsure to what extent real data necessitate this choice. The direct measure, while perhaps the simplest of the approaches, is susceptible to common method bias. In evaluating the PRISM using real data, the raw difference score and direct measure deviated the most from theory. The regression approach and analysis of partial variance were more in-step with theoretical expectations. We recommend researchers use either of those latter approaches. Fortunately, researchers already commonly use those approaches (e.g., Ahn & Kahlor, 2022; Ho et al., 2014; Hwang & Jeong, 2016; Kahlor et al., 2020; Kim & Lai, 2019; Yang & Kahlor, 2012), and they can take this conclusion as an assurance.

References

Ahn, J., & Kahlor, L. A. (2020). No regrets when it comes to your health: Anticipated regret, subjective norms, information insufficiency and intent to seek health information from multiple sources. *Health Communication, 35*(10), 1295–1302. doi:10.1080/10410236.2019.1626535

Ahn, J., & Kahlor, L. A. (2022). When experts offer conflicting information: A study of perceived ambiguity, information insufficiency, trustworthiness and risk information behaviors. *Health Communication, 38*(14), 3276–3286. doi:10.1080/10410236.2022.2146033

American Association for Public Opinion Research. (2023). *Standard definitions: Final dispositions of case codes and outcome rates for surveys*. Retrieved from https://aapor.org/wp-content/uploads/2024/03/Standards-Definitions-10th-edition.pdf

Census Bureau. (2020). *Age and sex composition in the United States: 2020*. Retrieved from https://www.census.gov/data/tables/2020/demo/age-and-sex/2020-age-sex-composition.html

Census Bureau. (2021a, April 21). *Educational attainment in the United States: 2020*. Retrieved from https://www.census.gov/data/tables/2020/demo/educational-attainment/cps-detailed-tables.html

Census Bureau. (2021b, September 14). *Income and poverty in the United States: 2020*. Retrieved from https://www.census.gov/library/publications/2021/demo/p60-273.html#:~:text=Median%20household%20income%20was%20%2467%2C521,median%20household%20income%20since%202011

Center for Climate Change Communication. (2022, July 20). *1. Global warming beliefs*. Retrieved from https://www.climatechangecommunication.org/all/climate-change-american-mind-april-2022-1/

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752–766. doi:10.1037/0022-3514.39.5.752

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York, NY: Guilford Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. doi:10.1080/10705510701301834

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York, NY: The Guilford Press.

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

Detenber, B., Rosenthal, S., Liao, Y., & Ho, S. S. (2016). Climate and sustainability| Audience segmentation for campaign design: Addressing climate change in Singapore. *International Journal of Communication, 10*, 4736–4758. Retrieved from https://ijoc.org/index.php/ijoc/article/view/4696

Dougherty, C. (2011). *Introduction to econometrics*. New York, NY: Oxford University Press.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich College Publishers.

Griffin, R. J., Dunwoody, S., & Neuwirth, K. (1999). Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. *Environmental Research, 80*(2), S230–S245. doi:10.1006/enrs.1998.3940

Griffin, R. J., Neuwirth, K., Dunwoody, S., & Giese, J. (2004). Information sufficiency and risk communication. *Media Psychology, 6*(1), 23–61. doi:10.1207/s1532785xmep0601_2

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2021). The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling. *Australasian Marketing Journal, 25*(1), 76–81. doi:10.1016/j.ausmj.2017.02.001

Ho, S. S., Detenber, B. H., Rosenthal, S., & Lee, E. W. J. (2014). Seeking information about climate change: Effects of media use in an extended PRISM. *Science Communication, 36*(3), 270–295. doi:10.1177/1075547013520238

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. doi:10.1080/10705519909540118

Huang, T., Leung, A. K.-y., Eom, K., & Tam, K.-P. (2022). Important to me and my society: How culture influences the roles of personal values and perceived group values in environmental engagements via collectivistic orientation. *Journal of Environmental Psychology, 80*, 101774. doi:10.1016/j.jenvp.2022.101774

Hunt, A., & Watkiss, P. (2011). Climate change impacts and adaptation in cities: A review of the literature. *Climatic Change, 104*(1), 13–49. doi:10.1007/s10584-010-9975-6

Hwang, Y., & Jeong, S.-H. (2016). Information insufficiency and information seeking: An experiment. *Science Communication, 38*(6), 679–698. doi:10.1177/1075547016673200

Kahlor, L. A. (2010). PRISM: A planned risk information seeking model. *Health Communication, 25*(4), 345–356. doi:10.1080/10410231003775172

Kahlor, L. A., Yang, J., Li, X., Wang, W., Olson, H. C., & Atkinson, L. (2020). Environmental risk (and benefit) information seeking intentions: The case of carbon capture and storage in Southeast Texas. *Environmental Communication, 14*(4), 555–572. doi:10.1080/17524032.2019.1699136

Kim, H. K., & Lai, C.-H. (2019). Seeking and sharing information about transboundary air pollution in Singapore: Effects of own and others' information insufficiency. *Environmental Communication, 14*(1), 68–81. doi:10.1080/17524032.2019.1597751

Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication Theory, 15*(2), 127–147. doi:10.1111/j.1468-2885.2005.tb00329.x

Li, J., Qin, P., Quan, Y., & Tan-Soo, J.-S. (2023). Using protection motivation theory to examine information-seeking behaviors on climate change. *Global Environmental Change, 81*, 102698. doi:10.1016/j.gloenvcha.2023.102698

Lu, H., Group, A. A., Winneg, K., Jamieson, K. H., & Albarracin, D. (2020). Intentions to seek information about the influenza vaccine: The role of informational subjective norms, anticipated and experienced affect, and information insufficiency among vaccinated and unvaccinated people. *Risk Analysis, 40*(10), 2040–2056. doi:10.1111/risa.13459

Mandloi, S. (2019, April 2). *Do Singaporeans really care about climate change?* Retrieved from https://www.eco-business.com/news/do-singaporeans-really-care-about-climate-change/

Rosenthal, S. (2011). *Personality and motivation in an augmented PRISM: Risk information seeking in the context of the indoor environment* (doctoral dissertation). The University of Texas at Austin, TX. Retrieved from https://repositories.lib.utexas.edu/bitstream/handle/2152/ETD-UT-2011-05-2906/ROSENTHAL-DISSERTATION.pdf?sequence=1&isAllowed=y

Rosenthal, S. (2013). Measuring differentials in communication research: Issues with multicollinearity in three methods. *Communication Methods and Measures, 7*(2), 106–125. doi:10.1080/19312458.2013.789837

Rosenthal, S. (2017). Structural equation modeling. In M. Allen (Ed.), *The SAGE encyclopedia of communication research methods* (Vol. 4, pp. 1683–1687). London, UK: SAGE Publications. doi:10.4135/9781483381411

Singapore Department of Statistics. (2021a). *Census of population 2020 statistical release 1*. Retrieved from https://www.singstat.gov.sg/-/media/files/publications/cop2020/sr1/cop2020sr1.pdf

Singapore Department of Statistics. (2021b). *Census of population 2020 statistical release 2*. Retrieved from https://www.singstat.gov.sg/-/media/files/publications/cop2020/sr2/cop2020sr2.ashx

Singapore Legal Advice. (2022, April 25). *Singapore fake news laws: Guide to POFMA (protection from online falsehoods and manipulation act)*. Retrieved from https://singaporelegaladvice.com/law-articles/singapore-fake-news-protection-online-falsehoods-manipulation/

Trumbo, C. W. (2002). Information processing and risk perception: An adaptation of the heuristic-systematic model. *Journal of Communication, 52*(2), 367–382. doi:10.1111/j.1460-2466.2002.tb02550.x

Wheeler, T., & von Braun, J. (2013). Climate change impacts on global food security. *Science, 341*(6145), 508–513. doi:10.1126/science.1239402

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Mason, OH: Sounth-Western, Cengage Learning.

Yang, Z. J., & Kahlor, L. A. (2012). What, me worry? The role of affect in information seeking and avoidance. *Science Communication, 35*(2), 189–212. doi:10.1177/1075547012441873

Yang, Z. J., Kahlor, L. A., & Li, H. (2013). A United States-China comparison of risk information–seeking intentions. *Communication Research, 41*(7), 935–960. doi:10.1177/0093650213479795

Appendix A: The Omitted Variable Bias in Raw Difference Score

Equation A1 represents a raw difference score and Equation A2 represents the regression approach. In these equations, *SI* indicates information-seeking intention, *ST* indicates information sufficiency threshold, *CK* indicates current knowledge, COV_i means the *i*th covariate, *rd* represents raw difference score approach, and *reg* indicates regression approach.

$$SI = B_{1rd}(ST - CK) + \sum_{i=3}^{n} B_i \cdot COV_i + B_0 \qquad (A1)$$

$$SI = B_{1reg} \cdot ST + B_2 \cdot CK + \sum_{i=3}^{n} B_i \cdot COV_i + B_0$$
$$= B_{1reg}(ST - CK) + (B_2 + B_{1reg})CK + \sum_{i=3}^{n} B_i \cdot COV_i + B_0 \qquad (A2)$$

Equation A1 does not incorporate current knowledge. Adding current knowledge as a covariate in Equation A1 (as seen in Equation A3), makes the model identical to Equation A2. Therefore, by adding *CK* as a covariate, the connection between insufficient information and the intention to seek information becomes the same as the regression approach.

$$SI = B_{1rd}(ST - CK) + B_{2rd} \cdot CK + \sum_{i=3}^{n} B_i \cdot COV_i + B_0$$
$$= B_{1rd} \cdot ST + (B_{2rd} - B_{1rd})CK + \sum_{i=3}^{n} B_i \cdot COV_i + B_0 \qquad (A3)$$

Table A1 displays the standardized coefficients of the raw difference score approach after including current knowledge as a covariate in the regression model. The results show that the relationship between information insufficiency and the intention to seek information remains positive in both countries regardless of the inclusion of other covariates.

Table A1. Standardized Coefficients of the Raw Difference Score Approach With Current Knowledge as a Covariate in OLS Regression.

Country	**Singapore**	**United States**	**Singapore**		**United States**	
Dependent variable	***SI***	***SI***	***SI***	***II***	***SI***	***II***
II	0.12^{***}	0.19^{***}	0.11^{***}		0.09^{***}	
CK	0.47^{***}	0.57^{***}	0.23^{***}	-0.57^{***}	0.23^{***}	-0.55^{***}
ATT			-0.04^{*}	0.25^{***}	-0.02	0.23^{***}
SN			0.50^{***}	-0.05^{**}	0.41^{***}	0.08^{**}

PC			0.04^{*}	-0.06^{**}	0.04^{*}	-0.11^{***}
NA			0.16^{***}	0.12^{***}	0.26^{***}	0.10^{***}
R^2	0.18	0.26	0.46	0.34	0.52	0.32

Note. SI = seeking intention. *II* = information insufficiency. *CK* = current knowledge. *ST* = information sufficiency threshold. *ATT* = attitude. *SN* = seeking-related subjective norms. *PC* = perceived seeking control. *NA* = negative affect. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.