Regular Submission

Organizational Research Methods
1–30

DOI: 10.1177/10944281211056520
journals.sagepub.com/home/orm
SAGE

The Quick and the Careless: The Construct Validity of Page Time as a Measure of Insufficient Effort Responding to Surveys

Nathan A. Bowling[1], Jason L. Huang[2], Cheyna K. Brower[1], and Caleb B. Bragg[3]

Abstract

Several recent studies have examined the prevention, causes, and consequences of insufficient effort responding (IER) to surveys. Scientific progress in this area, however, rests on the availability of construct-valid IER measures. In the current paper we describe the potential merits of the page time index, which is computed by counting the number of questionnaire pages to which a participant has responded more quickly than two seconds per item (see Huang et al., 2012). We conducted three studies (total $N = 1{,}056$) to examine the page time index's construct validity. Across these studies, we found that page time converged highly with other IER indices, that it was sensitive to an experimental manipulation warning participants to respond carefully, and that it predicted the extent to which participants were unable to recognize item content. We also found that page time's validity was superior to that of total completion time and that the two-seconds-per-item rule yielded a construct-valid page time score for items of various word lengths. Given its apparent validity, we provide practical recommendations for the use of the page time index.

Keywords

insufficient effort responding, careless responding, participant inattention, random responding, response effort

An alarming number of research participants display little effort when responding to survey questionnaires (Huang et al., 2012; Maniaci & Rogge, 2014; Meade & Craig, 2012). Some participants may, for instance, respond after hastily skimming questionnaire instructions,

[1]Department of Psychology, Wright State University, Dayton, Ohio, USA
[2]School of Human Resources and Labor Relations, Michigan State University, East Lansing, Michigan, USA
[3]Department of Psychological Science, Central Connecticut State University, New Britain, Connecticut, USA

Corresponding Author:
Nathan A. Bowling, Department of Psychology, Wright State University, Dayton, Ohio, USA.
Email: nathan.bowling@wright.edu

item stems, and response options. And in extreme cases, participants may respond to items without having read the questionnaire content at all. Such behavior, labeled as insufficient effort responding (IER; Huang et al., 2012), poses problems for users of questionnaire data. The presence of IER, for example, can undermine a basic researcher's ability to accurately test scientific models and it can undermine an applied researcher's ability to provide valid solutions to practical problems. Indeed, even small amounts of IER can produce misleading research findings (see Credé, 2010; Huang et al., 2012; Huang, Liu, et al., 2015; Schmitt & Stults, 1985).

Fortunately, researchers have made considerable progress in understanding the causes and prevention of IER (e.g., Bowling et al., 2016; Huang et al., 2012; Maniaci & Rogge, 2014). Further scientific progress, however, rests on the availability of flexible, construct-valid IER measures. Although research examining the measurement of IER has obvious implications for those who study the causes, consequences, and prevention of IER, it is also relevant for any organizational researcher or practitioner who uses self-report measures. Indeed, the effective measurement—and mitigation—of IER is applicable to several common data collection scenarios. Organizational practitioners who use crowdsourcing data as a means to establish scale norms, for instance, may find it necessary to omit responses provided high-IER participants. In such cases, screening for IER may produce observed norms that better reflect the scores found within an actual organizational sample (for a discussion of the systematic effects of IER on observed scale means, see Huang, Liu, et al., 2015). Similarly, measuring and preventing IER could help organizational practitioners make better informed decisions. This may be especially true in contexts where practitioners use low-stakes data, such as job incumbents' responses to a self-report personality assessment administered as part of a concurrent validation study, responses to an organizational climate survey, or responses to a self-report criterion measure used to assess the effectiveness of an organizational intervention.

As we argue below, the page time index (see Huang et al., 2012) has many desirable features that could make it a suitable measure of IER. In the following section we describe the page time index and explain its potential advantages. We then develop a nomological network for the IER construct. That network provided the basis for three studies (total $N = 1{,}056$) we conducted to test page time's construct validity.

Response Time Measures of IER

IER occurs when participants display "… low or little motivation to comply with survey instructions, correctly interpret item content, and provide accurate responses" (Huang et al., 2012, p. 100).[1] Although several measures have been used to assess IER (for reviews of various IER measurement methods, see Curran, 2016; DeSimone et al., 2015; Meade & Craig, 2012), the current research focuses on one measure: the page time index. As we describe below, the page time index is a type of response time measure.

Response time measures, which capture excessively fast responding, have been used in several studies to assess IER (e.g., Bowling et al., 2016; Huang et al., 2012; Maniaci & Rogge, 2014; Meade & Craig, 2012; Ward & Meade, 2018). The theoretical justification for this type of measure is straightforward: A minimal amount of time must elapse for participants to perform the steps that underlie effortful responding. These steps include reading and interpreting each item, recalling and integrating response-relevant information, and then translating that information into a response (see Krosnick, 1991; Tourangeau, 1984). Researchers may infer that a participant who completed a study questionnaire too quickly has bypassed one or more of these processes, and has thus engaged in IER.

Table 1. Common IER Detection Approaches, Benefits, and Limitations.

Detection Approach (Indices)	Description of Approach (Key citations)	Benefits	Limitations
Response Time (total completion time, page time)	Using improbably short response time to infer IER due to lack of attention and cognitive processing (Huang et al., 2012; Meade & Craig, 2012)	• Easily implemented • Unobtrusive, unlikely to lead to negative reactions from participants • Does not increase the burden placed on respondents • Unlikely to encounter resistance from stakeholders (e.g., management)	• Unable to capture distracted IER (e.g., turning attention to watch TV while responding carelessly) • Requires questionnaire platform to record time measures
Infrequency	Embedding items that most attentive participants will answer in the same direction (Huang et al., 2015; Meade & Craig, 2012)	• Easily implemented	• Requires the addition of special items, thus increasing questionnaire length • Possible resistance from stakeholders (e.g., management) • Potential negative reactions from participants without explanations • Can be misinterpreted by attentive respondents
Instructed Response	Embedding items that instruct participants to answer in a particular way (Kam & Chan, 2018)	• Easily implemented	• Requires the addition of special items, thus increasing questionnaire length • Possible resistance from stakeholders (e.g., management) • Potential negative reactions from participants without explanations • Sophisticated respondents may be on the lookout for these items
Inconsistency (individual reliability, psychometric synonyms, psychometric antonyms)	Computing within-person correlations based on two arrays that are expected to be strongly correlated (Huang et al.,	• Unobtrusive, unlikely to lead to negative reactions from participants • Requires no additional items	• Requires a large number of items (for psychometric antonyms/ synonyms) or scales (for individual reliability) to

(continued)

Table 1. (continued)

Detection Approach (Indices)	Description of Approach (Key citations)	Benefits	Limitations
	2012; Meade & Craig, 2012)	• Does not increase the burden placed on respondents • Unlikely to encounter resistance from stakeholders (e.g., management)	allow for the effective calculation of indices
Multivariate Outlier (Mahalanobis Distance)	Computing multivariate outlier statistic comparing individual responses on all items of a given measure against the sample's responses (Meade & Craig, 2012)	• Unobtrusive, unlikely to lead to negative reactions from participants • Requires no additional items • Does not increase the burden placed on respondents • Unlikely to encounter resistance from stakeholders (e.g., management)	• Requires a large number of items for a given measure to compute a single index
Long String (maximum long string, average long string)	Counting repeated endorsement of the same response option (Johnson, 2005; Meade & Craig, 2012)	• Unobtrusive, unlikely to lead to negative reactions from participants • Requires no additional items • Easily computed (see Yentes & Wilhelm, 2018) • Does not increase the burden placed on respondents • Unlikely to encounter resistance from stakeholders (e.g., management)	• Only assesses a particular pattern of IER • Requires similar or identical response scale throughout the survey (e.g., cannot change from 5-point to 7-point scale)
Self-report	Asking participants to report their level of attention upon completion of questionnaire (Huang et al., 2012; Meade & Craig, 2012)	• Easily implemented	• Requires the addition of special items, thus increasing questionnaire length • Assumes participants will pay attention to these items • Responses vulnerable to impression management

Response time measures have several desirable features relative to other IER indices (see Table 1 for a comparison of benefits and limitations across IER indices). First, they are easy to adapt to online questionnaires, as most popular electronic questionnaire platforms (e.g., Qualtrics, SurveyMonkey) provide a means to record participants' response times. Furthermore, response time measures do not require the addition of special items. This latter quality is desirable because the items used by some IER indices are obtrusive—and sometimes bizarre. Meade and Craig's (2012) infrequency scale, for instance, includes the items "I am paid biweekly by leprechauns" and "All my friends say I would make a great poodle." Some users are likely to resist the inclusion of such unusual content in their study questionnaires (e.g., organizational researchers may be hesitant to include such content in employee surveys). And by not requiring the inclusion of additional items, response time measures also help minimize the burden placed on respondents. This is important because excessive questionnaire length is a potential cause of IER (Bowling et al., 2021; Galesic & Bosnjak, 2009; Gibson & Bowling, 2020). In the following subsections we describe two kinds of response time measures—total completion time and page time. We then argue that relative to the former, the latter is a superior measure of IER.

Total Completion Time. Several studies have assessed IER using total completion time—the amount of time that elapses between when a participant begins and finishes a study questionnaire (e.g., Maniaci & Rogge, 2014; Meade & Craig, 2012; Ward & Pond, 2015). Prior research suggests that total completion time demonstrates poor convergent validity with other IER measures. Maniaci and Rogge (2014, Study 1), for instance, found that total completion time yielded correlations ranging from |.05| to |.13| with other IER indices. Although these correlations were in some instances statistically significant, they suggest minimal convergence. Similarly, Meade and Craig (2012, Study 1) found that total completion time loaded weakly onto all three IER factors they had identified in an exploratory factor analysis.

These findings raise two important questions: Why does total completion time—a measure with a seemingly strong theoretical basis—perform so poorly? And should researchers abandon response time indices? As we argue in the next subsection, the validity of total completion time may be undermined by the effects of nonresponding behavior. The page time index, by contrast, addresses this methodological challenge.

Page Time. The page time index, which was introduced by Huang et al. (2012), has been used to assess IER in several subsequent studies (e.g., Bowling et al., 2016; Gibson & Bowling, 2020; Ward & Meade, 2018). This index is computed from multi-page electronic questionnaires in which the researcher records the completion time of each participant for each questionnaire page. These data are then recoded according to whether or not the participant violated the two-seconds-per-item rule: If a participant completes a given questionnaire page at a rate faster than two seconds per item, then that participant is assumed to have responded carelessly to that page (recoded as "1"); however, if a participant completes a given questionnaire page at a rate equal to or slower than two seconds per item, then that participant is assumed to have responded carefully to that page (recoded as "0"). Averaging these recoded scores across the questionnaire pages produces the page time index.

Initial evidence suggests that the page time index correlates highly with other IER indices (see Bowling et al., 2016; Huang et al., 2012; Ward & Meade, 2018). Bowling et al. (2016, Study 5), for instance, found that page time yielded particularly high convergence with two popular IER indices—an infrequency index ($r = .78$, $p < .01$) and a psychometric antonym index ($r = .50$, $p < .01$). Such findings beg an important question: Why does page time appear to measure IER more effectively than does total completion time?

Page time may be a superior measure of IER because it is less susceptible to contamination from nonresponding behavior than is total completion time. Nonresponding behavior occurs when participants pause during questionnaire completion. Participants, for instance, may take self-imposed rest breaks or they may stop to attend to competing tasks (e.g., responding to text messages). Contamination of total completion time scores is possible because they reflect the participant's *typical pace* during questionnaire completion; thus, by engaging in nonresponding behavior, careless participants can obscure the fact that they have rushed through parts of the study questionnaire. Total completion time scores, therefore, may misclassify some careless participants as being careful (i.e., false negatives). Page time, on the other hand, is more resistant to contamination from nonresponding behavior due to the dichotomous scoring applied separately to each questionnaire page: Responding slowly on some pages will not negate violations of the two-seconds-per-item rule that have occurred on other pages.

Contamination from nonresponding behavior can also produce extreme total completion time values. Indeed, some participants may complete a questionnaire several hours—or even days—after they've begun responding (see Meade & Craig, 2012). The presence of extreme values undermines total completion time's construct validity, because participants with extremely large values are likely no more careful than are participants with more typical values. The page time index, however, is unlikely to yield extreme values. This is because page time values can only range from 0 (not being flagged on any survey page) to 1 (being flagged on all survey pages).

Nomological Network Used to Examine Page Time's Construct Validity

To examine a measure's construct validity, researchers must first develop a nomological network—a model that identifies a pattern of hypothesized relationships that exist between the focal measure and various external variables (see Cronbach & Meehl, 1955; Hinkin, 1998; Spector, 1991). Researchers can then infer support for the focal measure's construct validity to the extent that empirical evidence is consistent with the pattern of relationships predicted in the nomological network. As we describe in the following subsections, we adopted three approaches (see Murphy & Davidshofer, 2005) to comprehensively evaluate page time's construct validity. First, we evaluate page time's convergent validity with other measures of IER. Second, we examine whether page time is sensitive to an experimental manipulation that warned participants to respond carefully. Finally, we investigate page time's criterion-related validity in predicting participants' inability to recognize item content. By using these three approaches, the nomological network provided the basis for three studies we conducted to examine the construct validity of the page time index.

Other Measures of IER. Different measures of a given construct should converge with each other (Campbell & Fiske, 1959; Hinkin, 1998; Spector, 1991); thus, we expected page time to be positively correlated with other IER measures. Across three studies we examined page time's convergence with several other IER measures, including (a) a composite IER measure (except in Study 1, where we used an infrequency scale), (b) a long string index, and (c) self-reported carelessness. Because these IER indices have been reviewed and validated elsewhere (see Curran, 2016; DeSimone et al., 2015; Huang et al., 2012; Meade & Craig, 2012), we provide detailed descriptions of each IER index in the online supplemental materials.

Warning Participants to Respond Carefully. Researchers have argued that IER is largely a result of insufficient participant motivation (see Bowling et al., 2016; Huang et al., 2012; Meade & Craig, 2012). One means of increasing participant motivation—and thus a potential way of deterring IER—involves warning participants that engaging in IER will result in punishment (see Bowling et al., 2021; Gibson & Bowling, 2020; Huang et al., 2012). Huang et al., for instance, found that

participants who were warned that engaging in IER would result in the forfeiture of their research participation credits responded more carefully than did control participants. We thus predict that page time would be sensitive to the presence of a warning and would therefore yield lower scores among participants who have received a warning.

Inability to Recognize Item Content. Participants who engage in IER are unlikely to closely read and process the questionnaire's content. As a result, they may fail to later recognize item content from the study questionnaire. Thus, inability to recognize item content can serve as a criterion for IER, and we expect page time scores to positively predict the inability to recognize item content.

The Present Research

We conducted three studies to examine the construct validity of the page time index (see Table 2 for a summary of the features of these three studies). Specifically, we examined page time's convergence with other IER indices (Studies 1–3), its sensitivity to an experimental manipulation warning participants to respond carefully (Study 2), and its relationship with the inability to recognize item content (Studies 1 and 2). The three studies varied in length, thus offering us an opportunity to compare the validity of page time with that of total completion time across surveys of various lengths. As we explained earlier, we expected comparisons with total completion time to provide evidence of the page time index's superior validity. Finally, we tested the appropriateness of the two-seconds-per-item rule that is typically used to compute page time (see Bowling et al., 2016; Huang et al., 2012; Ward & Meade, 2018). Within all three studies, we compared the validity of page time scores computed using this rule with page time scores computed using various alternative computational rules (i.e., rules ranging from .6-s-per-item to 4.0-s-per-item). And in Study 3, we examined whether the two-seconds-per-item rule yielded a construct-valid page time score across items of various word lengths and whether participants' verbal ability influenced the convergent validity of the page time index.

Table 2. Summary of Features From Each Study.

	Study 1	Study 2	Study 3
Main survey length	107 items	223 items	325 items
Survey pages	5 pages	7 pages	16 pages
Average item length	5.31 words	10.72 words	6.25 words
Convergent validity	Infrequency scale	Infrequency	Infrequency
		Instructed Response Items	Individual Reliability
		Individual Reliability	Psychometric Synonyms
		Psychometric Synonyms	Psychometric Antonyms
		Psychometric Antonyms	Mahalanobis Distance
		Mahalanobis Distance	Long string
		Long string	
		Postsurvey Self-Report	
Criterion-related validity	Inability to recognize item content	Inability to recognize item content	
Response to manipulation		Warning	

Study 1 Method

Participants

Study 1 participants ($N = 197$) were employed undergraduate students recruited from a university in the Midwestern United States. Participants were asked about their work-related behaviors and perceptions, and they received research credits in exchange for their participation. Their average age was 21.26 years; 75% of participants were female; 61% were White, 20% were Black, and 11% were Asian American.

Measures

We administered the Study 1 questionnaire online. The main questionnaire comprised 107 items. Substantive measures include counterproductive work behavior, work centrality, satisfaction with supervisor, as well as items specifically developed for the current study (see online supplemental material for measures and sample items).

Page Time index. The Study 1 questionnaire was distributed across five pages (with 16 to 33 items on each page). We administered the questionnaire using Qualtrics, which allowed us to record the number of seconds participants spent responding to each page. We recoded the raw time data for each questionnaire page using the two-seconds-per-item rule: A participant was assumed to have responded carefully to a given page if he or she spent $\geq$2 s per item on that page (recoded as "0"); a participant was assumed to have responded carelessly to a given page if he or she spent <2 s per item on that page (recoded as "1"). We averaged these recoded values to create the page time index ($\alpha = .89$).

Total Completion Time index. We computed total completion time by summing the response times across the five survey pages ($M = 719$ s, $SD = 985$ s). Because total completion time scores display extreme positive skew, we computed the natural log of total completion time (Tabachnick et al., 2007). So that high scores would reflect high levels of IER, we reverse-scored the log-transformed scores by multiplying them by -1. In the remainder of this article, we refer to this reverse-scored transformed index as the *total completion time index*.[2]

IER Measure – Infrequency Scale. The relative brevity of Study 1 questionnaire made it difficult to compute reliable post-hoc IER indices. We thus used the infrequency approach to assess IER in Study 1 by embedding three items among the substantive survey items (Huang, Bowling et al., 2015). Attentive respondents were expected to respond in the same direction to an infrequency item—e.g., indicating disagreement to the statement "I work fourteen months in a year"—and failure to do so would indicate IER behavior. The three-item scale had a Cronbach's α of .70.

Inability to recognize item content. After completing the main survey, participants were informed that we were interested in ensuring their attention to the survey. They were asked to respond to 10 multiple choice questions about item content included in the main survey. For instance, an item in the main survey stated "I would stop working if I inherited a large sum of money from a family member." We assessed participants' recognition of the example embedded item using the following multiple-choice item:

Earlier in the questionnaire, we asked you whether you would stop working if you:

(a) *inherited a large sum of money from a family member*
(b) *became severely disabled in a work accident*

(c) *won a large sum of money in the lottery*
(d) *married a multimillionaire*

We coded participants' responses to each multiple-choice item based on whether they provided the correct response (coded "0") or one of the three incorrect responses (coded "1"). We then averaged these recoded values to produce an overall measure of their inability to recognize the item content. This multiple-choice assessment yielded an α of .84. We provide copies of both the embedded items and the multiple-choice items in the online supplemental materials.

Study 1 Results

Construct Validity of the Page Time Index

We examined page time's convergence with the infrequency IER scale and its criterion-related validity with the inability to recognize item content. We first present analyses in which we computed page time using the two-seconds-per-item rule; later we present analyses in which we computed page time using various alternative computational rules.

Convergence with Infrequency IER Scale. As shown in Table 3, the page time index yielded a significant positive relationship with the infrequency IER scale ($r = .56$, $p < .001$). This finding provides support for page time's convergent validity.

Relationship with Inability to Recognize Item Content. As shown in Table 3, the page time index was positively correlated with the inability to recognize item content ($r = .84$, $p < .001$). This provides support for page time's criterion-related validity.

Comparison of the Construct Validity of Page Time and Total Completion Time

In the previous section we described analysis examining page time's construct validity. We conducted similar analyses for the total completion time index, thus enabling us to compare page time's construct validity with that of total completion time.

Convergence with Infrequency IER Scale. The total completion time index yielded a significant positive relationship with the infrequency IER scale ($r = .41$, $p < .001$). Using z-test for dependent correlations (see Lee & Preacher, 2013; Steiger, 1980), we found that page time had a significantly stronger correlation with the infrequency IER scale ($z = 3.86$, $p < .001$) than did the total completion time index (see Table 3).

Table 3. Descriptive Statistics and Correlations for Study 1 Variables.

	M	SD	1	2	3	4
1. Page time index	.16	.30				
2. Total completion time	−6.23	.75	.70			
3. Infrequency scale	.17	.30	.56	.41	*.70*	
4. Inability to recognize item content	.19	.25	.84	.64	.59	.84
Steiger's Z			–	–	3.86	6.44

Note. $N = 197$. All $ps < .001$. Where appropriate, Cronbach's α appears in italics on the diagonal.
Total completion time is the natural log of total completion time, multiplied by −1 so that high scores indicate IER.
Steiger's Z: Z test comparing whether (1) *page time index* and (2) *total completion time* had significantly different associations with another variable.

Relationship with Inability to Recognize Item Content. The total completion time index yielded a significant ($r = .64$, $p < .001$) positive correlation with the inability to recognize item content. Page time, however, yielded a significantly stronger relationship with the inability to recognize item content than did the total completion time index ($z = 6.44$, $p < .001$; see Table 3).

Incremental Validity of Page Time Over Total Completion Time. Our next set of analyses considers information availability in practical survey administrations—if researchers have access to survey completion time on each survey page, then they will also have participants' total completion time values. Thus, we can ask two related questions: (a) Can researchers dichotomize the total completion time using the 2-s-per-item rule to better capture IER (rather than simply using the total completion time index)? and (b) Does page time offer incremental validity over the two indices that are based on total completion time (i.e., total completion time and dichotomized total completion time)?[3] We conducted hierarchical regression analyses to address these questions. As shown in Table 4, we entered the total completion time index in Step 1 of the hierarchical regression analysis, we added dichotomized total completion time (this index indicates whether or not a participant spent less than 2 s per item on average across the entire survey) in Step 2, and the page time index in Step 3.

When predicting the infrequency scale, the total completion time index accounted for 17% of variance (Step 1), while the dichotomized total completion time explained another 6% of variance (Step 2). In Step 3 of the analysis, the page time index accounted for an additional 10% of the variance in the infrequency scale. More importantly, neither the total completion time index nor the dichotomized total completion time could uniquely predict the infrequency scale once the effects of the page time index were controlled (see the Step 3 regression coefficients). The results were similar when predicting the inability to recognize item content: While the total completion time index accounted for 40% of variance in the criterion variable (Step 1), dichotomized total completion time explained additional 20% of variance (Step 2). In Step 3, the page time index contributed 11% of unique variance, while the other two predictors became non-significant (see the Step 3 regression coefficients).

Table 4. Incremental Validity of Page Time index in Study 1.

	Step 1 β	Step 2 β	Step 3 β	Relative weight	Rescaled relative weight
Outcome: Infrequency Scale					
1. Total completion time	.41***	.20*	.04	.07[a]	21%
2. Dichotomized total completion time		.31***	−.16	.08[a]	25%
3. Page time index			.67***	.17[b]	54%
ΔR^2	.17***	.06***	.10***		
Outcome: Inability to recognize item content					
1. Total completion time	.64***	.25***	.08	.15[a]	21%
2. Dichotomized total completion time		.59***	.07	.23[b]	32%
3. Page time index			.72***	.34[c]	47%
ΔR^2	.40***	.20***	.11***		

Note. $N = 197$. *$p < .05$; ***$p < .01$.
[a-d]Are used to indicate which relative weights are significantly different from each other.
Total completion time is the natural log of total completion time, multiplied by −1 so that high scores indicate IER.
For the same outcome variable, relative weights denoted by different superscripts are significantly different from each other (see Tonidandel & LeBreton, 2015).

Given the three time-based predictors were inherently correlated, we performed relative weight analysis (Johnson, 2000) to estimate the relative importance of each predictor. Specifically, using relative weights and rescaled relative weights, we assessed the amount of R^2 and the percentage of predicted variance attributable to each predictor (LeBreton et al., 2007). Furthermore, we estimated whether page time carried significantly stronger relative weight than the two indices based on total completion time using 10,000 bootstrapped replications (Tonidandel & LeBreton, 2015). The relative weight analysis matched the results from hierarchical regression: The page time index was the most important predictor for both infrequency scale and inability to recognize item content, and its contribution to their prediction significantly outweighed the other two indices.

The Construct Validity of Page Time Across Various Computational Rules

We compared the construct validity of page time scores computed using Huang et al.'s (2012) two-seconds-per-item rule with the construct validity of page time scores computed using alternative computational rules. Specifically, we computed page time scores using alternative computation rules at every .1-s interval ranging from .6-s-per-item to 4.0-s-per-item. We plotted the correlations between the resulting page time scores and (a) the infrequency IER scale and (b) the inability to recognize item content. As shown in Figure 1, the correlation between various page time scores and either (a) the infrequency IER scale or (b) the inability to recognize item content peaked near the two-seconds-per-item rule. These results suggest two-seconds-per-item might serve as a reasonable heuristic within applied survey projects.

Study 1 Discussion

In sum, Study 1 provided initial evidence of the page time index's construct validity. First, we found that page time converged with the infrequency IER scale. Second, in terms of criterion-related validity, the page time index strongly predicted the inability to recognize item content. Furthermore, we found that page time was more construct-valid than was total completion time. Finally, Study 1 also found support for the appropriateness of the two-seconds-per-item rule.

Study 1, however, is limited in two important ways. First, the relative brevity of the Study 1 questionnaire limited our ability to compute various post-hoc IER indices (see Huang, Bowling et al., 2015; Huang et al., 2012; Meade & Craig, 2012). The availability of multiple IER indices would provide a stronger test of page time's convergent validity. Second, the observational design used in Study 1 could not assess whether page time can detect different levels of IER under different survey instructions.

We conducted Study 2 as a constructive replication of Study 1, using a longer questionnaire that assessed different substantive content. Study 2 included a variety of IER indices, thus allowing for a more comprehensive evaluation of page time's convergent validity. In addition, following the idea that a valid measure of a construct should reflect different standings of respondents due to a manipulation on the intended construct (Murphy & Davidshofer, 2005), we included a warning (vs. control) manipulation in Study 2.

Study 2 Method

Participants

Study 2 participants ($N = 526$) were undergraduate students enrolled at a university in the Midwestern United States, which was different from the one used to recruit the Study 1 participants. Participants received research credits in exchange for completing the study questionnaire. Their

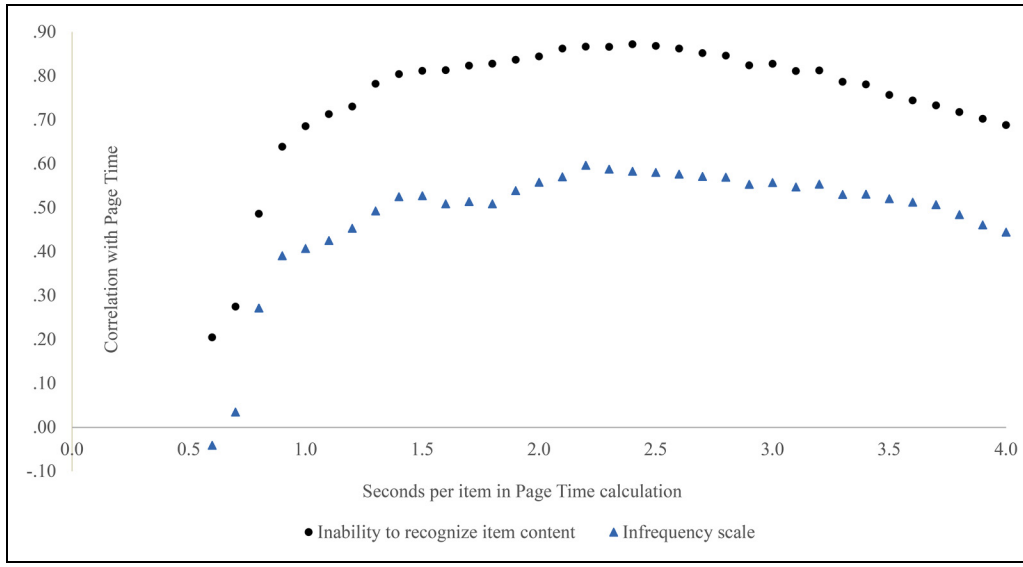

Figure 1. Correlation between page time based on different cutoffs and other IER indices in Study 1.

average age was 20.06 years; 60% of participants were female; 68% were White, 21% were Black, and 3% were Asian American.

Warning Manipulation

We randomly assigned each Study 2 participant to either a warning condition or to a control condition. Warning condition participants received a pre-questionnaire message stating that their levels of carelessness would be monitored and that they would forfeit their research participation credits if they were found to have responded carelessly. This warning was adapted from Huang et al. (2012). Control condition participants received no pre-questionnaire message. Participants received a debriefing statement upon completion of the study, and we awarded participants in both conditions full credit, regardless of whether or not they engaged in IER.

Measures

The Study 2 questionnaire was administered online. The main questionnaire comprised 223 items and included several substantive measures, such as five factor model personality traits, positive and negative affectivity, and life satisfaction (see online supplemental material for measures and sample items). These measures provided the means to compute several post hoc IER indices (e.g., the long string index; see Meade & Craig, 2012). We also embedded infrequency IER items and items that would enable us to later assess item content recognition (see below).

Page Time index. The main questionnaire was distributed across seven pages (with 20 to 38 items on each page). Similar to Study 1, we administered the Study 2 questionnaire using Qualtrics and used the two-seconds-per-item rule to compute the page time index ($\alpha = .92$).

Total Completion Time index. Because total completion time scores ($M = 2{,}356$, $SD = 20{,}924$ s) displayed extreme positive skewness, we computed the natural log of total completion time

(Tabachnick et al., 2007). As in Study 1, we operationalized total completion time by reverse-scoring (multiplying by −1) the natural log of the completion time scores.

Other IER Measures. We included the following IER indices in Study 2: (a) an infrequency scale ($\alpha = .85$); (b) an instructed-response index based on three items ($\alpha = .74$; e.g., "Please select *Very Inaccurate* for this question"); (c) individual reliability; (d) psychometric synonym index; (e) psychometric antonym index; (f) Mahalanobis Distance index; and (g) long string index (see below). Upon completion of the study, we also asked participants to self-report their IER. As most of these IER indices have been detailed in Huang et al. (2012) and Meade and Craig (2012), we describe only the long string index and self-reported IER here (see online supplemental material for approaches to calculate the other indices).

Long String IER index. We computed the maximum long string index, which has been used as a measure of IER within several previous studies (e.g., Bowling et al., 2021; Kam & Chan, 2018; Meade & Craig, 2012; Ward & Pond, 2015). To compute this index, we used a function from Yentes and Wilhelm's (2018) *R* package to identify the longest string of consecutive identical responses each participant provided on each of the seven questionnaire pages. The maximum long string index score for a given participant was the largest of these seven values.

Self-Reported IER index. We used Meade and Craig's (2012) nine-item diligence scale as a self-report measure of IER ($\alpha = .90$). This scale has been used in previous studies to assess IER (e.g., Francavilla et al., 2019; McKay et al., 2018). The diligence scale appeared at the end of the study questionnaire, immediately following the demographic items. A sample diligence item is "I carefully read every survey item." Each item was on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

Exploratory Factor Analysis for IER Indices. Given the large number of IER measures in Study 2, we conducted an exploratory factor analysis to examine the factor structure of the IER measures. Several decision rules, including the Kaiser criterion, the screeplot, parallel analysis (Hayton et al., 2004), and Velicer's (1976) MAP test, all indicated the presence of three factors. A direct oblimin rotation also indicated that three correlated factors were present: (a) a factor comprising most of the IER indices; (b) a factor comprising the long string index, and (c) a factor comprising the self-reported IER measure (see Table 5). Following previous studies (e.g., Bowling et al., 2016; Huang & DeSimone, 2021), we averaged the standardized scores of the six individual IER indices from the first factor to create a composite IER score ($\alpha = .80$). We used the maximum long string index and self-reported diligence as two additional IER measures.

Inability to Recognize Item Content. Similar to Study 1, the Study 2 participants completed a post-questionnaire assessment of their inability to recognize content from the study questionnaire. For this study, we embedded 11 particularly memorable items within the study questionnaire. An example item was "If my friends dared me to eat a live goldfish, I would probably do it." After they completed the study questionnaire, we informed participants that we were interested in assessing whether they paid attention to the study material and asked them to respond to 11 multiple-choice items. We coded participants' responses to each item based on whether they provided the correct response (coded "0") or one of the three incorrect responses (coded "1"). We then averaged these recoded values to produce an overall measure ($\alpha = .88$) of their inability to recognize the item content. We provide copies of both the embedded items and the multiple-choice items in the online supplemental materials.

Table 5. Factor Intercorrelations and Loadings From Exploratory Factor Analysis for IER Indices in Study 2.

	Factor 1 (Composite)	Factor 2 (Long String)	Factor 3 (Self-Report)
Factor 1	–		
Factor 2	.20	–	
Factor 3	.26	.36	–
Infrequency Scale	**.55**	.26	.26
Instructed Response Items	**.55**	.17	.21
Individual Reliability	**.76**	−.07	.01
Psychometric Synonyms	**.64**	−.09	.16
Psychometric Antonyms	**.60**	.08	.20
Mahalanobis Distance	**.47**	.00	−.26
Long string Max	−.04	**.94**	−.03
Long string Average	−.01	**.93**	−.02
Self-Report	.12	.00	**.88**
Use Me	.04	.06	**.54**

Note. $N = 526$. Principal axis factoring with oblimin rotation. The strongest loading for each index is bolded.

Study 2 Results

Construct Validity of the Page Time Index

We examined page time's convergence with other IER indices, its sensitivity to the warning manipulation, and its relationship with the inability to recognize item content. Together, these analyses provide a test of page time's construct validity. In the main analyses we operationalized page time using the two-seconds-per-item rule. Later in the Results section we present analyses in which we computed the page time index using computational rules ranging from .6-s-per-item to 4.0-s-per-item.

Convergence with Other IER Indices. As shown in Table 6, the page time index yielded significant ($p <$ .001) positive relationships with (a) the composite IER score ($r = .63$), (b) the maximum long string index ($r = .42$), and (c) the self-reported IER index ($r = .42$). These findings provide support for page time's convergent validity.

Sensitivity to the Warning Manipulation. We expected the page time index to be sensitive to the manipulation warning participants to respond carefully. As shown in Table 7, we found lower page time scores for participants assigned to the warning condition ($M = .02$) than for participants assigned to the control condition ($M = .06$; $t = 2.69$, $p < .01$; Cohen's $d = .23$). These results support page time's construct validity.

Relationship with Inability to Recognize Item Content. As shown in Table 6, the page time index was positively correlated with the inability to recognize item content ($r = .74$, $p < .001$). This provides evidence for page time's criterion-related validity.

Comparison of the Construct Validity of Page Time and Total Completion Time

In the previous section we described analysis examining page time's convergence with other IER indices, its sensitivity to a warning manipulation, and its relationship with the inability to recognize item content. We conducted similar analyses for total completion time index, thus allowing us to compare page time's construct validity with that of the total completion time index.

Table 6. Descriptive Statistics and Correlations for Study 2 Variables.

	M	SD	1	2	3	4	5	6	7
1. Warning condition	0.49	0.50	–						
2. Page time index	0.04	0.16	−.12**	–					
3. Total completion time	−7.06	0.52	−.06	.47***	–				
4. Composite IER	0.00	0.70	−.16***	.63***	.25***	*.80*			
5. Maximum long string index	6.51	4.61	−.04	.42***	.20***	.24***	–		
6. Self-reported diligence	1.90	0.91	−.22***	.42***	.17***	.45***	.30***	*.85*	
7. Inability to recognize item content	0.05	0.15	−.15***	.74***	.30***	.71***	.47***	.47***	*.88*
Steiger's Z			1.16	–	–	10.08***	5.27***	5.84***	12.33***

Note. $N = 526$. *$p < .05$; **$p < .01$; ***$p < .001$. Where appropriate, Cronbach's α appears on the diagonal.
For warning condition, Control Group = 0; Warning Group = 1. Total completion time is the natural log of total completion time, multiplied by −1 so that high scores indicate IER.
Steiger's Z: Z test comparing whether (2) *page time index* and (3) *total completion time* had significantly different associations with another variable.

Convergence with Other IER Indices. The total completion time index yielded significant (ps < .001) relationships with the composite IER score ($r = .25$), the long string index ($r = .20$), and the self-reported IER index ($r = .17$). A series of z-tests for dependent correlations found that page time had significantly ($p < .001$) stronger convergence with the composite IER score ($z = 10.08$), the long string index ($z = 5.27$), and the self-reported IER index ($z = 5.84$) than did the total completion time index (see Table 6), thus suggesting that the page time index has superior levels of construct validity.

Sensitivity to the Warning Manipulation. We examined whether the total completion time index was sensitive to the experimental warning manipulation. As shown in Table 7, the mean difference for the total completion time index was nonsignificant ($t = 1.47$, $p = .14$; Cohen's $d = .13$). However, a z-test comparing dependent correlations revealed that page time was not affected more strongly by the warning manipulation than was the total completion time index ($z = 1.16$, $p = .25$).

Relationship with Inability to Recognize Item Content. The total completion time index yielded a significant correlation ($r = .30$, $p < .001$) with the inability to recognize item content (see Table 6). A

Table 7. Results of t-Tests Examining the Effects of Warning Manipulation in Study 2.

	Control Condition (n = 267)		Warning Condition (n = 259)			
	M	SD	M	SD	*t*	Cohen's *d*
Page time	.06	.20	.02	.11	2.69**	.23
Total completion time	−7.03	.53	−7.09	.50	1.47	.13
Composite IER score	.11	.82	−.11	.53	3.63***	.32
Maximum long string index	6.69	4.87	6.32	4.33	.93	.08
Self-reported diligence	2.10	1.00	1.69	.76	5.28***	.46
Inability to recognize item content	.07	.18	.03	.10	3.47***	.30

Note. $N = 526$. *$p < .05$; **$p < .01$; ***$p < .001$.
Total completion time is the natural log of total completion time, multiplied by −1 so that high scores indicate IER.

follow-up z-test for dependent correlations found that page time yielded a significantly stronger relationship with the inability to recognize item content than did total completion time ($z = 12.33$, $p < .001$; see Table 6). This supports the superior construct validity of the page time index.

Incremental Validity of Page Time Over Total Completion Time. As in Study 1, we conducted hierarchical regression analysis to evaluate (a) whether dichotomizing total completion time produced better prediction of variables from the nomological network than did total completion time alone and (b) whether page time outperforms the two indices that are derived from total completion time. As shown in Table 8, these analyses yielded a consistent pattern across the four criterion variables: composite IER score, maximum long string index, self-reported diligence, and inability to recognize item content. Addressing the first question above, Step 2 of the analyses revealed that dichotomizing total completion time outperformed the total completion time index (ΔR^2 ranged from 5% to 24% across the four criteria), as the total completion time index was no longer associated with any criterion variable in the presence of dichotomized total completion time. Step 3 of these analyses suggest that the page time index is the only meaningful predictor of the outcomes (ΔR^2 ranged from 11% to 24% across the four criteria): While dichotomized total completion time remained a significant predictor, the negative regression coefficients indicate it served as a suppressor for the page time index. This conclusion is corroborated by the relative weight analysis, which we conducted in the same manner as in Study 1. Across the four outcomes, the page time index consistently contributed the most to the prediction, with rescaled relative weight greater than 60%, and it was significantly more important than the other two indices.

In short, Study 2 found that relative to the total completion time index, page time produced higher convergence with other IER indices, and was more strongly related to the inability to recognize item content. Although page time did not respond *more strongly* than the total completion time index to an experimental warning manipulation, tests of their sensitivity to the manipulation differed: Whereas page time was sensitive to the manipulation, the total completion time index was not. As a whole, these findings provide evidence for the superior validity of the page time index.

The Construct Validity of Page Time Across Various Computational Rules

To examine various cutoffs for page time, we computed page time scores using alternative computation rules at every .1-s interval ranging from .6-s-per-item to 4.0-s-per-item. We plotted the correlations between the resulting page time scores and (a) the composite IER score, (b) the long string index, (c) the self-reported IER index, and (d) the inability to recognize item content. As shown in Figure 2, page time scores computed using the two-seconds-per-item rule yielded reasonably strong correlations with these external variables; however, the correlations generally peaked at 1.5 s per item.

Study 2 Discussion

Findings from Study 2 corroborated those from Study 1 in showing page time's construct validity. The use of an experimental manipulation also provided further support for the validity of page time. The slight discrepancy in page time cutoffs between Study 1 and Study 2 (see Figures 1 and 2) is noteworthy. Aside from the motivational difference (i.e., half of participants in Study 2 received a warning), one key distinction is that Study 1 had shorter item stems (5.31 words per item on average) than Study 2 (10.72 words per item on average). It is possible that item length may influence the value of optimal page time cutoffs. As an item's length increases, the computational rule (i.e., number of seconds) that produced the most construct-valid page time score may also increase. Such an effect may occur simply because participants need more time to carefully read items of

Table 8. Incremental Validity of Page Time index in Study .

	Step 1 β	Step 2 β	Step 3 β	Relative weight	Rescaled relative weight
Outcome: Composite IER score					
1. Total completion time	.25***	.01	−.05	.02[a]	6%
2. Dichotomized total completion time		.51***	−.25***	.13[b]	30%
3. Page time index			.88***	.27[c]	64%
ΔR^2	.06***	.20***	.16***		
Outcome: Maximum long string index					
1. Total completion time	.20***	.07	.02	.02[a]	8%
2. Dichotomized total completion time		.26***	−.37***	.05[b]	26%
3. Page time index			.73***	.13[c]	66%
ΔR^2	.04***	.05***	.11***		
Outcome: Self-reported diligence					
1. Total completion time	.17***	.05	.00	.01[a]	6%
2. Dichotomized total completion time		.27***	−.37***	.05[b]	26%
3. Page time index			.75***	.14[c]	68%
ΔR^2	.03***	.06***	.12***		
Outcome: Inability to recognize item content					
1. Total completion time	.31***	.06	−.01	.04[a]	6%
2. Dichotomized total completion time		.55***	−.37***	.17[b]	29%
3. Page time index			1.07***	.37[c]	65%
ΔR^2	.09***	.24***	.24***		

Note. $N = 526$. ***$p < .001$.
[a-d]Are used to indicate which relative weights are significantly different from each other.
Total completion time is the natural log of total completion time, multiplied by −1 so that high scores indicate IER.
For the same outcome variable, relative weights denoted by different superscripts are significantly different from each other (see Tonidandel & LeBreton, 2015).

increasing length. Further, participants' verbal ability may also play a part in how page time functions (i.e., low-ability participants may require more time to carefully read questionnaire content). Given these possibilities, we conducted Study 3 to examine the potential moderating effects of (a) varying item lengths and (b) verbal ability.

Study 3 Method

Participants

Study 3 participants ($N = 333$) were undergraduate students enrolled at a university in the Northeastern United States. Participants received research credits in exchange for completing the study questionnaire. Their average age was 19.53 years; 63% of participants were female; 69% were White, 19% were Hispanic, and 15% were Black.

Item Length

We included three types of survey pages in the Study 3 questionnaire: (a) five pages consisting entirely of short items (average word length per item = 2.00 words), (b) five pages consisting entirely of medium-length items (average word length per item = 4.71 words), and (c) five pages consisting entirely of long items (average word length per item = 12.31 words). Each of these 15 pages included 20 items. We selected these items based on their item lengths rather than on the substantive constructs they assessed.

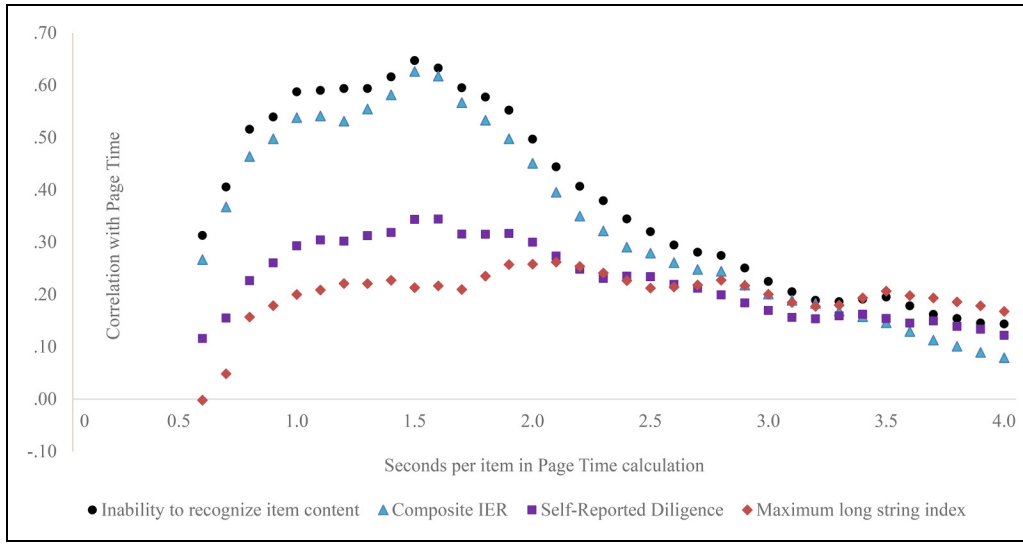

Figure 2. Correlation between page time based on different cutoffs and other IER indices in Study 2.

To minimize potential contrast effects stemming from switching back and forth between items of various lengths (e.g., abruptly switching from short to long items could have induced frustration among respondents), we asked participants to complete the five short-item pages first, the five medium-item pages second, and the five long-item pages last. We randomized the order of the pages within a given length type. All participants, for example, completed the five short-item pages first, but the specific order of those short-item pages was randomized across participants.

Measures

The Study 3 questionnaire was administered online. The main questionnaire comprised 325 items, including the 300 items described above, as well as the 20-item mini-IPIP measure (Donnellan et al., 2006) and five infrequency IER items (see online supplemental material for sample items). These measures provided both a medium in which to embed the IER items and a means of computing several post hoc IER indices.

Page Time index. The main questionnaire was distributed across 16 pages (with 20–25 items on each page). As in Studies 1 and 2, we used the two-seconds-per-item rule to compute the page time index ($\alpha = .93$). In the latter part of the Results section we present results using alternative cutoffs for page time.

Total Completion Time index. We computed the total completion time index with the same method used in Studies 1 and 2. One total completion time score ($M = 1{,}847$, $SD = 1{,}116$ s) was computed for each participant, rather than separately for the shot item, medium-length item, and long item pages. As in Studies 1 and 2, we operationalized the total completion time index by reverse scoring (multiplying by -1) the natural log transformed total completion time scores.

Other IER Measures. Study 3 included the following IER indices: (a) an infrequency scale based on five items ($\alpha = .83$); (b) individual reliability; (c) psychometric synonym index; (d) psychometric antonym index; (e) Mahalanobis Distance index; and (f) long string index (both maximum and average; see Study 2). Rather than computing them separately for the short, medium-length, and long item pages, we computed each IER index once for each participant.

Exploratory Factor Analysis for IER Indices. Similar to Study 2, we performed exploratory factor analysis on the various IER indices in Study 3. All decision rules, including the Kaiser criterion, the screeplot, parallel analysis Hayton et al. (2004) and Velicer's (1976) MAP test suggested a two-factor solution. As seen in Table 9, the two correlated factors, extracted through direct oblimin rotation, reflected a composite IER dimension and long string dimension. Although the infrequency scale loaded unexpectedly on the long string dimension, a separate exploratory factor analysis without the two long string indices yielded a single-factor solution. Thus, consistent with Study 2 (also see Bowling et al., 2016; Huang & DeSimone, 2021), we computed an IER composite by averaging the standardized scores of the five IER indices ($\alpha = .75$).

Verbal Ability. Prior to the survey, we assessed participants' verbal ability using the 10-item vocabulary test from General Social Survey (Cor et al., 2012). Each item presents participants with a target word and asked them to identify, from a list of five choices, one option that most closely shares the meaning of the target word.

Study 3 Results

Construct Validity of the Page Time Index

Study 3 examined page time's convergence with two other IER measures—the composite IER score and the maximum long string. We conduct the main analyses using the two-seconds-per-item rule. Later we compute page time using various alternative computational rules. As shown in Table 10, the page time index yielded a significant positive relationship with both the composite IER score ($r = .67$, $p < .001$) and the long string index ($r = .50$, $p < .001$). These findings provide support for page time's convergent validity.

Comparison of the Construct Validity of Page Time and Total Completion Time

As in Studies 1 and 2, we compared page time's validity with that of the total completion time index. Total completion time was positively related to both the composite IER score ($r = .30$, $p < .001$) and the long string index ($r = .20$, $p < .001$). Using z-tests for dependent correlations (Steiger, 1980), we found that page time had stronger correlations with both the composite IER score ($z = 8.94$, $p < .001$) and the long string index ($z = 6.54$, $p < .001$) than did the total completion time index (see Table 10). These findings suggest that compared to total completion time, page time provides superior construct validity.

Table 9. Factor Intercorrelations and Loadings From Exploratory Factor Analysis for IER Indices in Study 3.

	Factor 1 (Composite)	Factor 2 (Long String)
Factor 1	—	
Factor 2	.35	—
Infrequency Scale	.30	**.55**
Individual Reliability	**.75**	.01
Psychometric Synonyms	**.76**	.12
Psychometric Antonyms	**.63**	.00
Mahalanobis Distance	**.44**	−.05
Long string Max	−.07	**.99**
Long string Average	−.05	**.89**

Note. $N = 333$. Principal axis factoring with oblimin rotation. The strongest loading for each index is bolded.

Table 10. Descriptive Statistics and Correlations for Study 3 Variables.

	M	SD	1	2	3	4	5
1. Verbal ability	0.49	0.15					
2. Page time index	0.09	0.20	−.19***				
3. Total completion time	−7.39	0.50	−.16**	.59***			
4. Composite IER	0.00	0.71	−.21***	.67***	.30***	.75	
5. Maximum long string index	7.25	4.65	−.07	.50***	.20***	.40***	
Steiger's Z			—	—	—	8.94***	6.54***

Note. $N = 333$. *$p < .05$; **$p < .01$; ***$p < .001$. Where appropriate, Cronbach's α appears on the diagonal.
Steiger's Z: Z test comparing whether (2) *page time index* and (3) *total completion time* had significantly different associations with another variable.

Incremental Validity of Page Time Over Total Completion Time. As in Studies 1 and 2, we further examined the incremental validity of dichotomized total completion time and the page time index using hierarchical regression analyses (see Table 11). As shown in Step 2, dichotomized total completion time added to the prediction of the composite IER score ($\Delta R^2 = 12\%$), but not to the prediction of the maximum long string index ($\Delta R^2 = 0\%$) beyond the effects of the total completion time index. In Step 3, the page time index contributed to the prediction of both outcomes ($\Delta R^2 = 25\%$ for the composite IER score and 29% for the maximum long string index). With the presence of the page time index, the other two indices based on total completion time were no longer meaningfully associated with the criterion variables, and the negative significant coefficients suggest they merely served as suppressors. Again, relative weight analysis revealed a similar pattern of results: The page time index was more important than both total completion time and dichotomized total completion time for predicting either outcome, accounting for more than 70% of predicted variance.

The Construct Validity of Page Time Across Various Computational Rules

We compared the convergent validity of page time scores computed using alternative computational rules, ranging from .6 s per item to 4.0 s per item (see Figure 3). Across item

Table 11. Incremental Validity of Page Time index in Study 3.

	Step 1 β	Step 2 β	Step 3 β	Relative weight	Rescaled relative weight
Outcome: Composite IER score					
1. Total completion time	.30***	.12*	−.14**	.04[a]	9%
2. Dichotomized total completion time		.39***	.01	.09[b]	20%
3. Page time index			.75***	.33[c]	71%
ΔR^2	.09***	.12***	.25***		
Outcome: Maximum long string index					
1. Total completion time	.20***	.17**	−.11*	.03[a]	9%
2. Dichotomized total completion time		.06	−.35***	.04[a]	13%
3. Page time index			.81***	.26[b]	78%
ΔR^2	.04***	.00	.29***		

Note. $N = 333$. *$p < .05$; **$p < .01$; ***$p < .001$.
[a-d]Are used to indicate which relative weights are significantly different from each other.
Total completion time is the natural log of total completion time, multiplied by −1 so that high scores indicate IER.
For the same outcome variable, relative weights denoted by different superscripts are significantly different from each other (see Tonidandel & LeBreton, 2015).

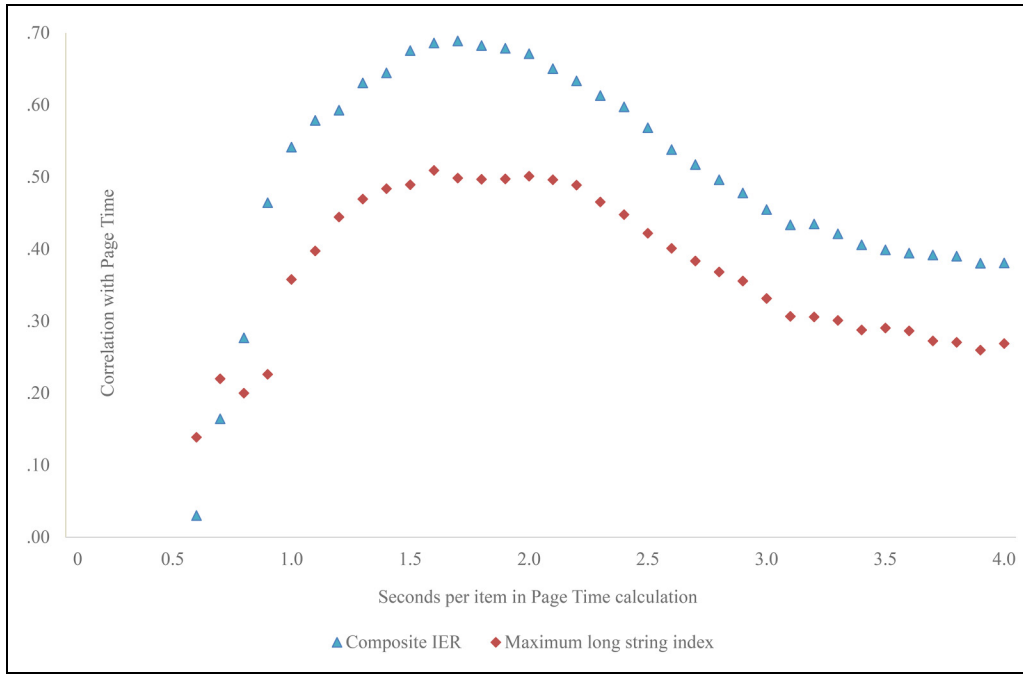

Figure 3. Correlation between page time based on different cutoffs and other IER indices in Study 3.

lengths, the 2.0-s-per-item rule appeared to perform as well as, if not better than, other alternative cutoffs.

Next, we plotted three sets of results—one set for the five pages containing short items (Figure 4), one for the five pages containing medium-length items (Figure 5), and one for the five pages

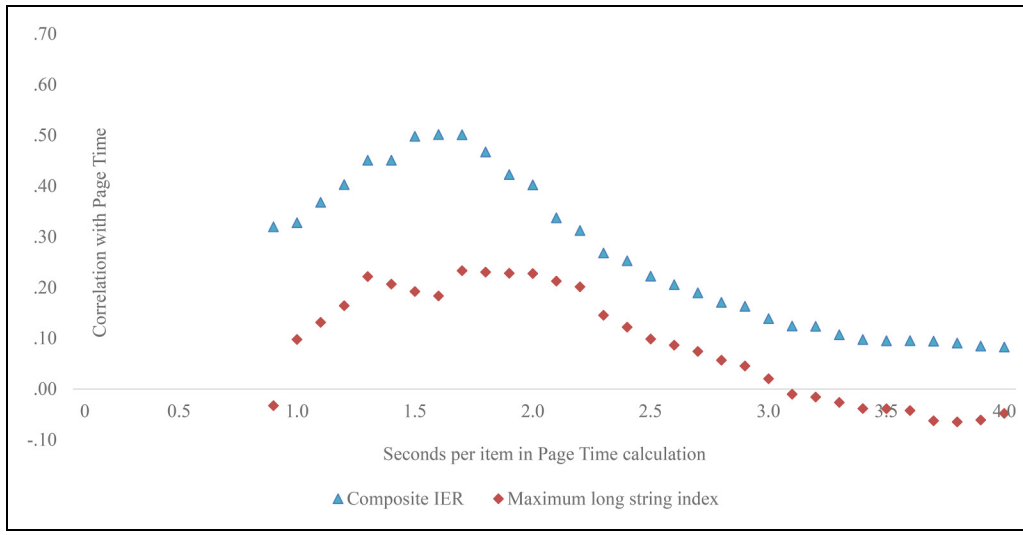

Figure 4. Correlation between page time for short items based on different cutoffs and other IER indices in Study 3. *Note*. Page time index did not flag any IER cases when cutoff was below .9 s per item.

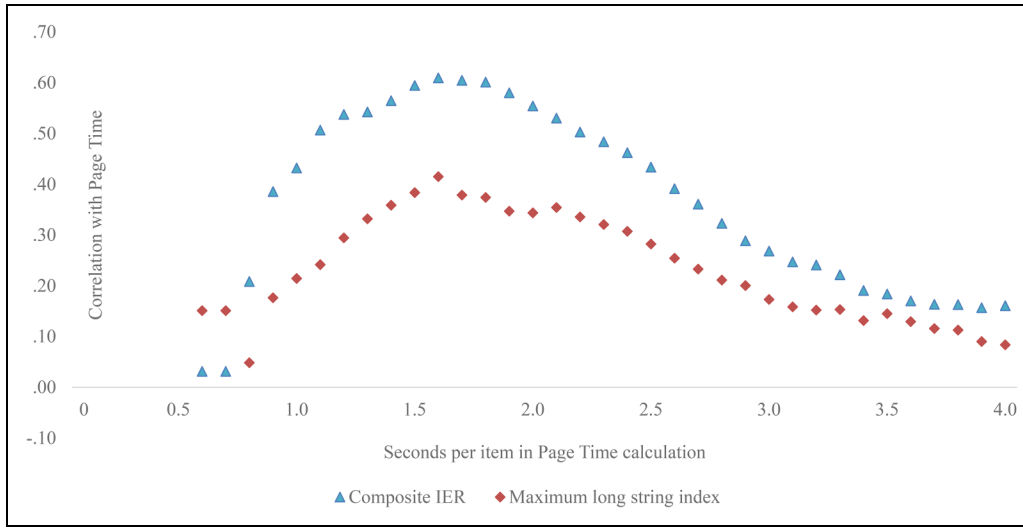

Figure 5. Correlation between page time for medium items based on different cutoffs and other IER indices in Study 3.

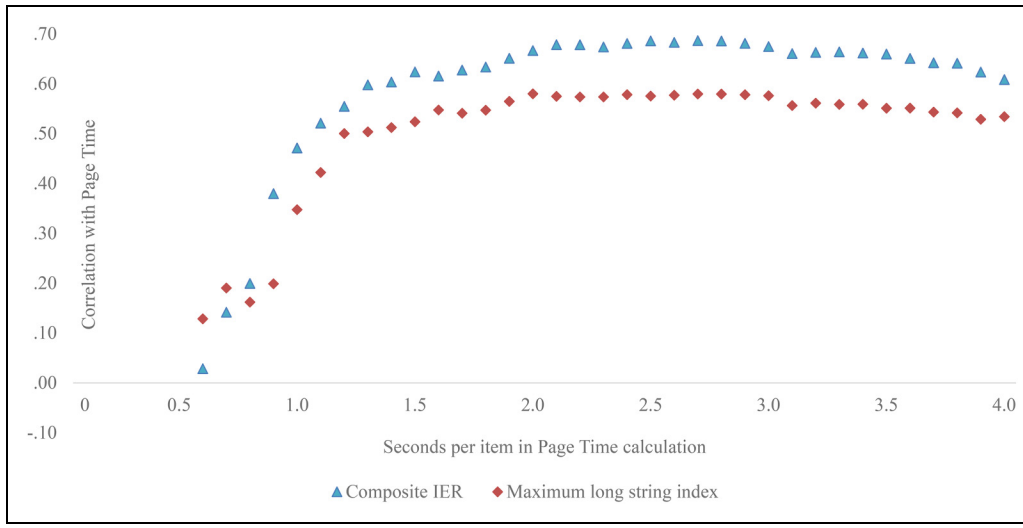

Figure 6. Correlation between page time for long items based on different cutoffs and other IER indices in Study 3.

containing long items (Figure 6). These analyses suggest that a 1.5-s-per-item rule may provide maximal validity for questionnaires that comprise short or medium-length items, but that a 2.0-s-per-item rule may provide maximal validity for questionnaires that comprise long items.

Moderating Role of Verbal Ability

We performed moderated regression analyses to examine whether verbal ability moderated the convergent validity of page time. Results in Table 12 indicate that verbal ability did not interact with

Table 12. Moderated Regression Examining Verbal Ability as a Potential Moderator in Study 3.

	Model 1		Model 2		Model 3	
Predictors for models	B	β	B	β	B	β
Outcome: Composite IER						
1. Page time index	2.37***	.67***	2.31***	.66***	2.34***	.66***
2. Verbal ability			−.39*	−.09*	−.39*	−.08*
3. Page time × verbal ability					.57	.03
R^2	.45***	.46***	.46***			
ΔR^2	.45***	.01*	.00			
Outcome: Maximum long string index						
1. Page time index	11.59***	.50***	11.73***	.51***	12.25***	.53***
2. Verbal ability			.91	.03	.98	.03
3. Page time × verbal ability					8.66	.06
R^2	.25***	.25***	.26***			
ΔR^2	.25***	.00	.00			

Note. $N = 333$. *$p < .05$; **$p < .01$; ***$p < .001$.

page time to predict either the composite IER or the long string index. The page time index, therefore, appears to be equally valid across respondents with different levels of verbal ability.

Study 3 Discussion

Study 3 provided a replication and extension of Studies 1 and 2. First, we found that the page time index converged well with other measures of IER and that it outperformed total completion time and dichotomized total completion time. We also found that although the two-seconds-per-item rule generally provided an acceptable means of computing the page time index; however, this rule might require adjustment depending on item length. Finally, we found that page time generally performed well as a measure of IER across participants with varying levels of verbal ability.

General Discussion

We used three independent samples to examine page time's construct validity as a measure of IER. Our results found that page time yielded a pattern of relationships that was consistent with the pattern predicted by the nomological network. Specifically, the page time index (a) generally converged with other IER indices, (b) was sensitive to an experimental manipulation that warned participants to respond carefully, and (c) was related to the inability to recognize item content. These findings strongly support the construct validity of the page time index. Similar analyses for the total completion time index, however, suggest that it is a relatively less effective measure of IER. Indeed, results of hierarchical regression and relative weight analysis both suggest that reliance on total completion time, even after dichotomizing it using the 2-s-per-item rule, would lead to a failure to capture meaningful variance in IER.[4] It is noteworthy that we observed consistent results across the three datasets, given recent concerns over the failure of many research findings to replicate (see Open Science Collaboration, 2015; Pashler & Wagenmakers, 2012).

We also found support for the use of Huang et al.'s (2012) two-seconds-per-item rule as the basis for computing page time: Analyses from all three of our studies suggest that page time scores computed using the Huang et al., rule generally produced correlations with external variables that were either similar to or stronger than the correlations we observed for page time scores computed using

alternative computational rules, such as a one-second-per-item rule or a four-seconds-per-item rule. In Study 3, however, we found that page time might yield slightly higher construct validity among shorter items when a 1.5-s-per-item rule was used. As a whole, these findings are noteworthy because although the two-seconds-per-item rule has been used in several studies as the basis for computing page time scores (e.g., Gibson & Bowling, 2020; Huang et al., 2012; Ward & Meade, 2018), researchers had not previously examined the effectiveness of this rule relative to that of other computational rules. It is also of note that the two-seconds-per-item rule appears to be a reasonably effective means of computing page time scores for items of various word lengths (see Study 3).

Implications

As we noted in the Introduction section, the page time index has many desirable qualities. First, it can be easily adapted to online questionnaires. This is possible because several electronic questionnaire platforms (e.g., Qualtrics) allow researchers to record participants' response times. Given the popularity of online questionnaires, this is a particularly important feature. Furthermore, the page time index, unlike other IER indices (e.g., the infrequency index; see Meade & Craig, 2012), does not require the addition of special questionnaire items. This latter feature allows researchers to unobtrusively assess IER while minimizing the burden placed on participants. Furthermore, compared to IER detection based on total completion time, the page time approach represents a more reliable and precise measure of IER. These features, when combined with the validity evidence reported in the current paper, make the page time index a particularly appealing IER measure.

Although the current research focused on the construct validity of the page time index, we also found evidence supporting the construct validity of other IER indices. The composite IER index, for instance, yielded a pattern of results that largely matched the pattern predicted by the nomological network (e.g., it was sensitive to the warning manipulation and it predicted participants' inability to recognize item content; see Study 2). As such, the composite IER index also appears to be effective measures of IER.

Practical Advice for Using the Page Time Index

Having provided evidence for the construct validity of the page time index, it is important for us to discuss how researchers can most effectively use the index to improve the quality of their research. We structure this discussion around two primary methods researchers can use to combat IER: (a) data screening and (b) prevention.

Data Screening. The page time index could be used to identify which respondents have engaged in excessive levels of IER. Data from respondents who have been flagged by the page time index could then be excluded from further analyses. In addition to the computational rule used to compute page time (e.g., 2-s-per-item rule vs. some alternative rule), researchers need to determine the value for page time cut scores—that is, for how many pages should a participant have to be flagged to qualify for removal from the dataset? The operationalization of page time may vary based on several factors, such as the nature and motivation level of the sample (e.g., students, job incumbents, online samples), questionnaire length, and questionnaire content. When possible, we recommend that researchers conduct pilot research to identify how to best compute page time in their main study. Researchers, for example, may ask highly motivated participants to respond to the pilot questionnaire either attentively or carelessly. Such a design would help identify the computational rule and precise cut score for page time that distinguish participants from the two conditions (see Huang et al., 2012). Another approach for pilot research is to instruct highly motivated respondents to complete the questionnaire as quickly as possible without compromising the quality of their responses. By examining page time distributions of attentive yet speedy respondents, this approach

would establish the computation rule and a priori cut score for page time. We strongly recommend the a priori identification of cut score values, because the use of cut score values identified post hoc can increase the risk of capitalizing on chance.

When pilot data are unavailable, researchers may follow the present study's findings to use the 2-s-per-item computational rule for page time and then examine the frequency distribution of the page time index to determine an appropriate cutoff value. Visual inspection of the frequency distribution of page time scores may be helpful for identifying an appropriate cut score, particularly when that distribution is bimodal (see Wise, 2019). In Study 1, for example, while 71% of respondents did not violate the 2-s-per-item rule on any pages. The remaining respondents' page time values resembled a bimodal distribution. Excluding the high end of the distribution would remove 28 of the 197 respondents (14%). We urge researchers to determine appropriate screening criteria before conducting any substantive analyses.

While our studies examined page time computation for typical organizational survey items (see online supplemental material), the computation of page time is likely different for scenario/vignette studies (Aguinis & Bradley, 2014), where reading becomes a much heavier component relative to self-reflection and response generation. For these studies, researchers may consider the typical reading speed of the sample of interest and screen respondents who respond unusually quickly (see Huang, 2014). Reading speed can be used in conjunction with manipulation checks that assess accurate recognition of scenario details.

Aside from removing the most egregious IER cases, researchers may also decide to remove partial data from a suspected IER case based on page time. That is, if someone rushed through a single survey page, then his or her responses on that particular page may be removed. Regardless of the decision rule used to screen data, we urge researchers to select a data screening rule *before* testing any substantive hypotheses.

IER Prevention. Despite its potential merits, researchers should be cautious about using the data screening approach to combat IER, because omitting data from high-IER participants (a) can harm statistical power (i.e., due to a reduction in sample size) and (b) can undermine the representative of a given sample, since a particular type of person is most likely to engage in IER (e.g., respondents who are low in either conscientiousness, agreeableness, or emotional stability; see Bowling et al., 2016). It is thus often preferable to prevent IER from occurring. The page time index is useful within this latter preventative approach. Low page time index values, for instance, can be used to document that IER is generally absent from a given dataset or that a method intended to prevent IER (e.g., a warning manipulation) has been effective.

Regardless of which approach a researcher uses, it is important to consider whether the page time index should be used alone or in combination with one or more other IER indices. There are good reasons to expect that page time is effective as a stand-alone measure. First, page time is strongly related to other IER indices (see Tables 3, 6, and 10); as a result, it would generally agree with other IER indices regarding which particular participants have or have not engaged in IER. Second, page time can be used in situations where various limitations preclude the use of other IER indices (for a discussion of the limitations of various IER indices, see Table 1). Page time, for instance, doesn't require the addition of special items, thus minimizing the burden placed on research participants. Contrast this with the infrequency, instructed-response, and self-reported approaches, which do require the addition of special items. Furthermore, the page time index can be used to assess IER in relatively short questionnaires (see Study 1). It is thus more desirable than the inconsistency approaches, which require the identification of a sufficient number of item pairs or scale halves, thus generally preventing their use within short questionnaires.

Furthermore, page time provides a non-intrusive means of assessing IER. Questionnaire participants, in other words, would generally be unaware that researchers are assessing their response time.

This feature would be important if researchers were concerned that knowledge that IER was being assessed would change participants' responses. Finally, the page time index may be considered an omnibus measure of IER. Spending insufficient time responding to questionnaire items, in other words, may capture most forms of IER—including IER that manifests as a "random" or alternatively as a systematic response pattern. It may be that most participants who engage in IER do so as a means of conserving time, thus making page time a relatively direct measure of IER. Other IER indices, in contrast, may be more limited in the breadth of IER behavior they assess. The long string approach, for example, assesses a very specific form of systematic IER—providing the same response to consecutive items—and the infrequency approach assesses "random" forms of IER (see Meade & Craig, 2012).

Limitations and Future Research

We note a few limitations of the current research. First, each of our three studies used a student sample. Although we expect that the page time index would perform effectively within other types of samples, research is needed to directly test this assumption. Second, it is unclear whether our findings regarding the effectiveness of various page time computational rules can be generalized to questionnaires that include other types of substantive content. Each of our study questionnaires consisted of relatively brief self-report items (see Table 2). Although we found consistent support for the effectiveness of the two-seconds-per-item rule, we encourage researchers to consider the content of their study questionnaires before choosing a particular computational rule. Two seconds, for instance, may allow sufficient time to respond carefully to a simple one-word item (e.g., items from Goldberg's (1992) Unipolar Big Five Markers), but it may allow too little time to respond carefully to a lengthy, complex item (e.g., items from James's (1998) conditional reasoning test). Similarly, the computation rules used to compute page time scores within vignette studies would likely need to be modified to accommodate the reading-heavy demands that such research places on participants. Future studies, therefore, should examine the validity of the page time index within questionnaires that contain various types of item content, including vignette studies.

A third limitation pertains to the context in which page time is applicable—when the time to respond to a questionnaire is unconstrained. While we believe most organizational survey research falls under this context, exceptions do exist, such as when participants are allotted a short amount of time to complete a survey before it times out. Responding to a survey under time constraint may change how the page time index performs: Short response time on a given page could reflect either an honest attempt to complete the survey as quickly as possible or IER behavior. In such circumstances, it could even be possible that total completion time may outperform page time as an indicator of IER.

Finally, we should note that the page time index detects IER by identifying excessively fast responding; however, it is conceivable that some IER behavior may manifest in slower response times and thus would be overlooked by the page time index. A distracted respondent, for instance, may spend some time watching TV before endorsing response options carelessly. However, it is reasonable to expect the preponderance of IER behavior would manifest in overly fast response time *on some survey pages*. That is, IER behavior likely stems from a lack of motivation and effort to read, interpret, and react to survey materials, which should translate into minimal time spent responding to survey pages. Page time, therefore, may provide an omnibus and relatively direct measure of IER, in contrast to other IER indices (e.g., infrequency, inconsistency, long string indices) that cannot directly capture this speeding behavior. Nevertheless, more research is needed to assess situations in which page time fails to capture IER. Such research may ask participants to use a verbal protocol approach to recall their specific response behavior when engaging in IER.

Summary

An alarming number of participants respond carelessly to survey questionnaires (Huang et al., 2012; Maniaci & Rogge, 2014; Meade & Craig, 2012). As a result, researchers have recently given increased attention to the study of IER. Further progress in this area, however, rests on the availability of valid IER measures. In the current paper we used three independent samples to examine the construct validity of one such IER measure—the page time index. As a whole, our findings suggest that the page time index is a valid measure of IER. We also found that Huang et al.'s two-seconds-per-item rule is generally an appropriate means of computing the page time index. As a result of our findings, we encourage researchers to use the page time index to assess IER.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Nathan A. Bowling https://orcid.org/0000-0003-3768-4939
Jason L. Huang https://orcid.org/0000-0001-9548-0542

Supplemental material

Supplemental material for this article is available online.

Notes

1. This behavior has been referred to using several terms, including "insufficient effort responding" (Huang et al., 2012), "careless responding" (Meade & Craig, 2012), "participant inattention" (Maniaci & Rogge, 2014), and "random responding" (Credé, 2010). Because of its recent popularity, we use the term "insufficient effort responding."
2. Across three studies, untransformed total completion time showed weaker associations with all other variables than the current transformed index. Results based on untransformed total completion time are available from the first author.
3. We thank two anonymous reviewers for suggesting that we probe these research questions.
4. Supplemental regression and relative weight analyses that included only total completion time and the page time index (i.e., dichotomized total completion time was omitted) indicated even stronger evidence for the greater relative importance of the page time index. As shown in the tables included at the end of the supplemental file, we observed this pattern across all three studies. It may be worth noting that dichotomizing total completion time using the 2-seconds-per-item rule accounted for some of the variance attributed to the page time index.

References

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351-371. https://doi.org/10.1177/1094428114547952

Bowling, N. A., Gibson, A. M., Houpt, J. W., & Brower, C. K. (2021). Will the questions ever end? Within-person increases in careless responding during questionnaire completion. *Organizational Research Methods*, *24*(4), 718-738. https://doi.org/10.1177/1094428120947794

Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. *Journal of Personality and Social Psychology*, *111*(2), 218-229. https://doi.org/10.1037/pspp0000085

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

Cor, M. K., Haertel, E., Krosnick, J. A., & Malhotra, N. (2012). Improving ability measurement in surveys by following the principles of IRT: The Wordsum vocabulary test in the General Social Survey. *Social Science Research*, *41*(5), 1003-1016. https://doi.org/10.1016/j.ssresearch.2012.05.007

Credé, M. (2010). Random responding as a threat to the validity of effect size estimates in correlational research. *Educational and Psychological Measurement*, *70*(4), 596-612. https://doi.org/10.1177/0013164410366686

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. https://doi.org/10.1037/h0040957

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4-19. https://doi.org/10.1016/j.jesp.2015.07.006

DeSimone, J. A., Harms, P. D., & DeSimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior*, *36*(2), 171-181. https://doi.org/10.1002/job.1962

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment*, *18*(2), 192-203. https://doi.org/10.1037/1040-3590.18.2.192

Francavilla, N. M., Meade, A. W., & Young, A. L. (2019). Social interaction and internet-based surveys: Examining the effects of virtual and in-person proctors on careless response. *Applied Psychology*, *68*(2), 223-249. https://doi.org/10.1111/apps.12159

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349-360. https://doi.org/10.1093/poq/nfp031

Gibson, A. M., & Bowling, N. A. (2020). The effects of questionnaire length and behavioral consequences on careless responding. *European Journal of Psychological Assessment*, *36*(2), 410-420. https://doi.org/10.1027/1015-5759/a000526

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*(1), 26-42. https://doi.org/10.1037/1040-3590.4.1.26

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191-205. https://doi.org/10.1177/1094428104263675

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104-121. https://doi.org/10.1177/109442819800100106

Huang, J. L. (2014). Does cleanliness influence moral judgments? Response effort moderates the effect of cleanliness priming on moral judgment. *Frontiers in Psychology*, *5*(1276), 1-8. https://doi.org/10.3389/fpsyg.2014.01276

Huang, J. L., Bowling, N. A., Liu, M., & Li, Y. (2015). Detecting insufficient effort responding with an infrequency scale: Evaluating validity and participant reactions. *Journal of Business and Psychology*, *30*, 299-311. https://doi.org/10.1007/s10869-014-9357-6

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, *27*(5), 99-114. https://doi.org/10.1007/s10869-011-9231-8

Huang, J. L., & DeSimone, J. A. (2021). Insufficient effort responding as a potential confound between survey measures and objective tests. *Journal of Business and Psychology*, *36*(5), 807-828. https://doi.org/10.1007/s10869-020-09707-2

Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology*, *100*(3), 828-845. https://doi.org/10.1037/a0038510

James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods*, *1*(2), 131-163. https://doi.org/10.1177/109442819812001

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality*, *39*(1), 103-129. https://doi.org/10.1016/j.jrp.2004.09.009

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, *35*(1), 1-19. https://doi.org/10.1207/S15327906MBR3501_1

Kam, C. C. S., & Chan, G. H. H. (2018). Examination of the validity of instructed response items in identifying careless respondents. *Personality and Individual Differences*, *129*, 83-87. https://doi.org/10.1016/j.paid.2018.03.022

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213-236. https://doi.org/10.1002/acp.2350050305

LeBreton, J. M., Hargis, M. B., Griepentrog, B., Oswald, F. L., & Ployhart, R. E. (2007). A multidimensional approach for evaluating variables in organizational research and practice. *Personnel Psychology*, *60*(2), 475-498. https://doi.org/10.1111/j.1744-6570.2007.00080.x

Lee, I. A., & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, *48*, 61-83. https://doi.org/10.1016/j.jrp.2013.09.008

McKay, A. S., Garcia, D. M., Clapper, J. P., & Shultz, K. S. (2018). The attentive and the careless: Examining the relationship between benevolent and malevolent personality traits with careless responding in online surveys. *Computers in Human Behavior*, *84*, 295-303. https://doi.org/10.1016/j.chb.2018.03.007

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437-455. https://doi.org/10.1037/a0028085

Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (6th ed.). Pearson Prentice Hall.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science (new York, N Y)*, *349*(6251), 1-8. https://doi.org/10.1126/science.aac4716

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science a crisis of confidence? *Perspectives on Psychological Science*, *7*(6), 528-530. https://doi.org/10.1177/1745691612465253

Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, *9*(4), 367-373. https://doi.org/10.1177/014662168500900405

Spector, P. E. (1991). *Summated rating scale construction: An introduction*. Sage.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245-251. https://doi.org/10.1037/0033-2909.87.2.245

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics*. Pearson.

Tonidandel, S., & LeBreton, J. M. (2015). RWA Web: A free, comprehensive, web-based, and user-friendly tool for relative weight analyses. *Journal of Business and Psychology*, *30*, 207-216. https://doi.org/10.1007/s10869-014-9351-z

Tourangeau, R. (1984). Cognitive science and survey methods. In *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines* (Vol. 15, pp. 73-100).

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321-327. https://doi.org/10.1007/BF02293557

Ward, M. K., & Meade, A. W. (2018). Applying social psychology to prevent careless responding during online surveys. *Applied Psychology: An International Review*, *67*(2), 231-263. https://doi.org/10.1111/apps.12118

Ward, M. K., & Pond, S. B. (2015). Using virtual presence and survey instructions to minimize careless responding on Internet-based surveys. *Computers in Human Behavior*, *48*, 554-568. https://doi.org/10.1016/j.chb.2015.01.070

Wise, S. L. (2019). An information-based approach to identifying rapid-guessing thresholds. *Applied Measurement in Education*, *32*(4), 325-336. https://doi.org/10.1080/08957347.2019.1660350

Yentes, R. D., & Wilhelm, F. (2018). careless: Procedures for computing indices of careless responding. *R package version*, *1*, 2018.

Author Biographies

Nathan A. Bowling (Ph.D. 2005, Central Michigan University) is a professor of psychology at Wright State University and a fellow of both the Society for Industrial and Organizational Psychology and the Association for Psychological Science. He has published over 60 peer-reviewed journal articles, with most of his research focusing on occupational health psychology, counterproductive work behavior, and research participant carelessness. As a PI, Nathan has been awarded nearly $1M in external funding since 2013. He also serves as a research consultant to the National Institute for Occupational Safety and Health and is an associate editor for the *Journal of Occupational Health Psychology*.

Jason L. Huang is an Associate Professor in the School of Human Resources and Labor Relations at Michigan State University. He received his Ph.D. in Organizational Psychology from Michigan State University. His research focuses on individual differences, employee adaptability, training and transfer, and response effort. Jason is currently an associate editor for *Journal of Management* and serves on the editorial boards for *Journal of Applied Psychology*, *Journal of Vocational Behavior*, and *Journal of Business Psychology*.

Cheyna K Brower is a personnel psychologist in the Research Team of the National Center for Organizational Development, Department of Veterans Affairs. She is a graduate of Auburn University and received her Ph.D. from Wright State University in 2020. Cheyna has studied various methods to detect careless responding on surveys. Her recent work includes developing assessments about employee attitudes, the impact of COVID-19 pandemic on the workplace, executive team development and followers' perceptions, and the impact of telework on the workplace.

Caleb B. Bragg is an Associate Professor of Psychology at Central Connecticut State University. He earned his Ph.D. in 2015 from Wright State University. Caleb has published articles in top peer reviewed journal, with most of his work focusing on counterproductive work behavior, research participant carelessness, and how personality variables (including self-control) impact the workplace and college classrooms.