The Leadership Quarterly 35 (2024) 101809

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

The research transparency index

Herman Aguinis [a], Zhuyi Angelina Li [b,*], Maw Der Foo [c]

[a] School of Business, The George Washington University, United States
[b] Renmin Business School, Renmin University of China, China
[c] Nanyang Business School, Nanyang Technological University, Singapore

ARTICLE INFO

Keywords:
Transparency
Research transparency index
Research transparency tool
Reproducibility
Replicability

ABSTRACT

Research transparency is critical for credible and trustworthy theory and subsequent practices and policymaking. However, checking for transparency is a laborious and time-consuming task. To facilitate this process, we introduce the *Research Transparency Index* (RTI v. 1.0). The program, available at https://www.hermanaguinis.com, enables users to assess the level of transparency in both unpublished and published manuscripts. RTI provides feedback on the transparency of manuscripts describing quantitative research across key research stages: theory, design, measurement, data analysis, and reporting of results. RTI (a) assists authors in enhancing the transparency of their manuscripts before submitting them to journals and conferences, (b) provides students with guidelines to improve their understanding of research transparency, and (c) provides reviewers and journal editors with a tool to assess manuscripts and offer developmental feedback to authors. RTI saves authors, students, reviewers, and editors time by providing an automated assessment of transparency criteria, which can be updated in the future, given that we make the Python code available. Also, it promotes a culture of transparency, fostering trust and credibility in the scholarly community and among users of the knowledge we produce (e.g., organization and policy decision-makers).

Introduction

The recent scandal concerning research misconduct that involved possible fabrications and refusal to share data by researchers across several universities has garnered significant attention (Gelles, 2023). This and other incidents underscored the critical role of research transparency, which instills confidence in the trustworthiness of results and can aid in correcting false positive findings (O'Boyle et al., 2017; Simmons et al., 2011) and reducing the occurrence of biased outcomes (Banks et al., 2016). Specifically, methodological transparency refers to "the degree of detail and disclosure about the specific steps, decisions, and judgment calls made during a scientific study" (Aguinis et al., 2018, p.83). Thus, transparency is crucial in producing credible and trustworthy research, leading to clearer contributions to theory and improved practices and policymaking (Aguinis, 2025).

Debate is ongoing in management and adjacent fields regarding the extent to which reproducibility (i.e., do conclusions converge when a different team of researchers analyzes the same data?) and replicability (i.e., do conclusions converge when a different team of researchers conducts a similar study?) are similarly applicable to quantitative and qualitative research. However, there is agreement that transparency is a *sine qua non* characteristic of high-quality and high-impact scholarship, encompassing both quantitative and qualitative approaches (Aguinis, 2025; Aguinis et al., 2017; Aguinis & Solarino, 2019; Banks et al., 2016; Wulff et al., 2023).

The Leadership Quarterly (*LQ*) has introduced a novel approach by implementing formal and systematic methodological checks to identify and address common empirical deficiencies before publishing articles (Friedrich et al., 2009; Wulff et al., 2023). However, assessing transparency can be time-consuming, especially when involving multiple transparency issues. For this reason, expecting reviewers and journal editors to examine all aspects thoroughly is unrealistic. Accordingly, to alleviate the burden of manual review and assist authors and editors, we introduce the *Research Transparency Index* (RTI v. 1.0). Specifically, RTI is a free tool available at https://www.hermanaguinis.com that allows users to assess the level of transparency in both unpublished and published manuscripts. While transparency is essential for qualitative and quantitative research, we initially designed this tool for quantitative

* Corresponding author at: Renmin Business School, Room 733 No. 59 Zhongguancun Street, Beijing, #100872, China.
E-mail addresses: haguinis@gwu.edu (H. Aguinis), lizhuyi@rmbs.ruc.edu.cn (Z.A. Li), mawderfoo@ntu.edu.sg (M. Der Foo).
URL: https://orcid.org/0000-0002-3485-9484 (H. Aguinis), https://orcid.org/0000-0003-1815-7017 (Z.A. Li), https://orcid.org/0000-0002-4887-081X (M. Der Foo).

https://doi.org/10.1016/j.leaqua.2024.101809
Received 20 October 2023; Received in revised form 7 June 2024; Accepted 26 June 2024
Available online 2 July 2024

Table 1
Criteria for Assessing Transparency in Quantitative Research (Italicized Criteria Are Checked by the Research Transparency Index–RTI). (see Aczel et al., 2020, Aguinis et al., 2018, and Wulff et al., 2023 for a review of these criteria).

Research Stage	Transparency Criterion
Theory	1. *Research strategy (e.g., inductive, deductive, abductive)*
	2. Level of theory, measurement, and analysis (e.g., individual, dyadic, organizational)
	3. *Supported and unsupported hypotheses for post-hoc hypotheses with detailed explanations, if any*
Research Design	
	4. *Type of research design (e.g., passive observation, experimental)*
	5. *Data collection procedure (e.g., surveys, interviews)*
	6. *Location of data collection (e.g., countries)*
	7. *Sampling method (e.g., purposeful, snowball, convenience)*
	8. Sample characteristics (e.g., students versus full-time employees; employment status, hierarchical level in the organization; sex; age; race)
	9. Sample size and explain its appropriateness
	10. If a power analysis was conducted before initiating the study or after the study's completion, report the results and explain if and how they affect the interpretation of the study's results
	11. If common method variance was addressed, state the theoretical rationale (e.g., failure to correlate with other self-report variables) and study design (e.g., temporal separation and use of self- and other-report measures) or statistical remedies (e.g., Harman one-factor analysis) used to address it
	12. Explanation of which control variables were included and which were excluded and why, along with how they influenced the variables of interest
Measurement	
	13. Report all measures used and provide a conceptual definition of the construct
	14. Provide evidence of construct validity (e.g., correlation tables including all variables, results of item and factor analysis)
	15. *If scales were altered, report how and why (e.g., dropped items, changes in item referent). Provide psychometric evidence regarding the altered scales (e.g., criterion-related validity)*
	16. *Report the exact items used in every reversed scale, if any*
	17. *If scores are aggregated, report measurement variability within and between units of analysis (e.g., $r_{wg(j)}$ and ICC)*
	18. *If range restriction was assessed, specify the type of range restriction and provide details regarding the rationale for the decision to correct or not correct*
	19. If constructs were translated, report the translation procedure (e.g., direct translation, back translation); explain whether there were cross-cultural issues associated with data collection and how these issues were resolved
Data Analysis	
	20. *Specific analytical method used and why it was chosen (e.g., EFA vs. CFA)*
	21. *Software used and its version*
	22. *If tests for outliers, anomalies, or other inconsistencies were conducted, report details such as methods and decision rules used to identify outliers and anomalies; steps (if any) taken to manage outliers and anomalies (e.g., deletion, Winsorization, transformation); the rationale for those steps; and results with and without outliers and anomalies*
Reporting of Results	
	23. *Results of missing-data analysis (e.g., sensitivity analysis); method (e.g., imputation, deletion) used to address missing data; information (even if speculative) as to why missing data occurred*
	24. *Exact response rate*
	25. Results of all tests of assumptions associated with the analytical method (e.g., normality, heteroscedasticity, independence, covariance among levels of repeated measures, homogeneity of treatment-difference variances, and group size differences in ANOVA)
	26. Complete descriptive statistics (e.g., mean, standard deviation, maximum, minimum) for all variables, as well as correlation and (when appropriate) covariance matrices
	27. *Effect size estimates, CI for point estimates, and information used to compute effect sizes (e.g., within-group variances, degrees of freedom of statistical tests); identify precise estimate used when referring to "effect size" (e.g., Cohen's d, r, R^2)*
	28. *Exact p-values to two decimal places; do not report p-values compared to cut-offs (e.g., $p < .05$ or $p < .01$); and authors should not report only asterisks to signal p-value thresholds*
	29. *Identify both unstandardized and standardized coefficients*
	30. Report any preregistered study designs, analysis plans, and hypotheses on a recognized platform before data collection begins; report any alterations to preregistered plans transparently
	31. Report and provide open access to all raw data used for analyses, except when restricted by privacy concerns

Note. $r_{wg(j)}$ is an index used to determine group inter-rater agreement, computed by comparing the actual group variance to the expected random variance. ICC refers to the intraclass correlation coefficient, a reliability estimate of the consistency of measurements made by different raters. EFA: exploratory factor analysis. CFA: confirmatory factor analysis. Cohen's *d* is an effect-size estimate showing the standardized difference between two group means. Pearson's *r* refers to Pearson's correlation coefficient, which examines a linear relation between two variables. R^2 refers to the coefficient of determination used to examine the proportion of variance explained by the independent variables in an outcome variable.

approaches, given the existence of consensual guidelines.

Our automated approach is not entirely new, as many journals already use tools such as iThenticate, Turnitin, Copyscape, PlagScan, Unicheck, and Quetext as part of the manuscript submission process. RTI operates in a comparable automated manner, but instead of assessing a manuscript's originality, it provides users with quantitative and qualitative information on methodological transparency across key research stages: (a) theory, (b) design, (c) measurement, (d) data analysis, and (e) reporting of results. Therefore, RTI aims to alleviate the burden of manual review, support the transparency assessment by authors and editors, and enhance the efficiency and effectiveness of the manuscript submission process. RTI is a valuable tool that can be used for any submission, whether to a conference or a journal. Of course, authors would still need to ensure that any submission meets the standards that particular organizations and journals set forth by reviewing author submission guidelines, editorials, and policies provided.

The remainder of our article is organized as follows. First, we discuss transparency criteria. Second, we describe how RTI is useful for the several stakeholders involved in producing and consuming scientific knowledge: Authors, students, reviewers and journal editors, as well as organization and policy decision-makers. Third, we describe the RTI and illustrate its use with two manuscripts. Fourth, we detail two studies we conducted to gather evidence regarding the validity of the RTI. Finally, we offer additional recommendations on how to use the RTI, including caveats and suggestions for future research on transparency.

Transparency Criteria Across Five Research Stages

The first research stage, theory, encompasses several aspects, including the level of theory, the type of analysis, and the specification of a priori hypotheses and their distinction from post-hoc hypotheses (Aguinis et al., 2018; Wulff et al., 2023). This stage is crucial for

Table 2
Recommendations for How Multiple Stakeholders Can Use and Benefit from the Research Transparency Index (RTI).

Users	Recommendations
Authors	• Based on the transparency index, authors know how transparent their manuscripts are. • The transparency report helps authors identify whether they have missed critical research details before they submit the manuscript for conference presentation and journal publication. • Authors can refer to the transparency report for suggestions and examples on improving their document's transparency index.
Students	• Given the transparency report, RTI is a learning tool that helps students improve paper transparency. • Students can better understand how to write a transparent paper using RTI to check for inappropriate reporting styles and missing details. • RTI can assist supervisors in coaching their students regarding general approaches to transparent academic writing. For example, supervisors can use RTI in transparency improvement workshops to demonstrate how students can enhance the transparency of their manuscripts.
Reviewers and editors	• Reviewers and editors can use RTI to help them quickly review papers regarding transparency issues. • A transparency report will be shown, including suggestions on whether manuscripts have severe transparency issues. • Reviewers and editors can refer to the report to determine whether manuscripts can be considered for further review or must be returned to authors for revision before formal submission.
Organization and policy decision-makers	• By showing the transparency level, the RTI report can help organizations and policy decision-makers select transparent published articles as sources for creating policies and making decisions. • Even with little knowledge in a specific field, organizations and policymakers can utilize RTI to understand whether specific published studies they may want to rely on are sufficiently transparent.

enhancing inferential reproducibility and drawing consistent inferences (Kraimer et al., 2023). The second stage, design, involves factors such as the type of design, data collection procedures, sampling methods, power analysis, common method variance, and control variables (Aguinis et al., 2018; Antonakis, 2017; Wulff et al., 2023). The third stage, measurement, focuses on, for example, clear construct definitions, psychometric properties related to reliability and validity, statistics justifying aggregation, and consideration of range restrictions (Aguinis et al., 2017, 2018, 2020; Wulff et al., 2023). The fourth stage, data analysis, involves reporting the details of various techniques, such as the chosen analytical method, software, syntax, and tests for outliers and possible data anomalies (Aguinis et al., 2018; Wulff et al., 2023). Lastly, the reporting stage pertains to how researchers present their results and conclusions (Aguinis et al., 2018; Banks et al., 2016; Toth et al., 2020). This stage includes reporting issues such as missing data analysis, assumption tests, descriptive statistics, effect sizes, and *p*-values.

As summarized in Table 1, each stage comprises numerous criteria, resulting in 31. It can be time-consuming for authors and reviewers to examine each criterion thoroughly. Furthermore, due to conscious and unconscious biases, authors may unintentionally overlook crucial research details while immersed in the writing process. For instance, when rounding off *p*-values, some authors may report cutoffs (e.g., $p < .05$) instead of providing exact *p*-values (Bettis, 2012). A survey conducted among active researchers revealed that approximately 11 % (95 % CI: 9 %–13 %) of respondents had rounded off *p*-values to get below a cutoff score (Banks et al., 2016). Although we now know the importance of reporting exact *p*-values (Antonakis, 2017), verifying each reported *p*-value and identifying inadvertent errors is laborious.

RTI Users: Authors, Students, Reviewers and Editors, and Organization and Policy Decision-Makers

Table 2 summarizes how multiple stakeholders can use RTI in the research production and dissemination process. Upon analyzing a manuscript with the RTI, users indicate whether they are (a) authors and students or (b) reviewers and journal editors. RTI then generates tailored transparency reports for each user group. For authors, RTI generates a transparency index for each criterion and offers suggestions to assist in identifying any research details that may be missing before submitting their manuscript for publication (or conference) consideration. This feedback aids authors in enhancing the transparency of their research and improving their understanding of how to make their work more transparent. Second, students will find RTI valuable as a learning tool. Using RTI to identify inappropriate reporting styles and missing details, students can better understand and learn how to write a more transparent paper. So, RTI serves as a self-learning resource that provides students with guidelines to comprehend transparency issues they need to address. RTI can also support supervisors in coaching students on general approaches to achieve transparent academic writing. For reviewers and editors, RTI simplifies the review process for submitted journal papers by identifying transparency issues. RTI generates a transparency report accompanied by developmental feedback on whether the submitted manuscripts exhibit serious transparency issues and if they may need to be returned to the authors. This feature saves reviewers and editors valuable time and effort, reducing the need to manually check for missed details every time a new manuscript is submitted. Finally, RTI can be used by organization and policy decision-makers interested in learning about the relative transparency of published research to decide what knowledge to use as input for their practices and policies.

Using the RTI

We developed RTI using Python 3, incorporating the *pdfminer* natural language processing package. This tool utilizes text mining techniques, including text matching and text clustering, based on the latest advancements in data analytics. To use the RTI, available at https://www.hermanaguinis.com (including the source code), users download a self-extracting zipped folder with executable files designed for Mac OS and Windows systems. We designed it specifically so RTI runs locally on users' computers, and manuscripts are not uploaded anywhere externally—and, therefore, not shared with anyone else. RTI v. 1.0 is tailored explicitly for quantitative papers and does not provide transparency-related suggestions for qualitative papers, conceptual papers, papers with multiple studies, or papers utilizing experimental designs and secondary data.[1]

The process of using RTI is as follows. The top panel of Fig. 1 illustrates the interface, which requires all papers (unpublished or published) to be submitted in PDF format. The next step involves completing a disclosure form, as depicted in the bottom panel of Fig. 1. To further improve transparency and address the issue that certain criteria cannot

[1] The decision to exclude experimental designs and studies using secondary data in this initial version of RTI was driven by the fact that experimental designs typically involve fewer variables and fewer data to report, resulting in fewer transparency concerns compared to other research designs that involve a larger number of variables and data collected by the authors themselves (Hill et al., 2022). Relatedly, qualitative papers have a distinct replication logic (Knight et al., 2022), which may require different transparency tools or be addressed in future versions of RTI. As transparency norms evolve, future versions can include other research designs and features (e.g., qualitative studies).

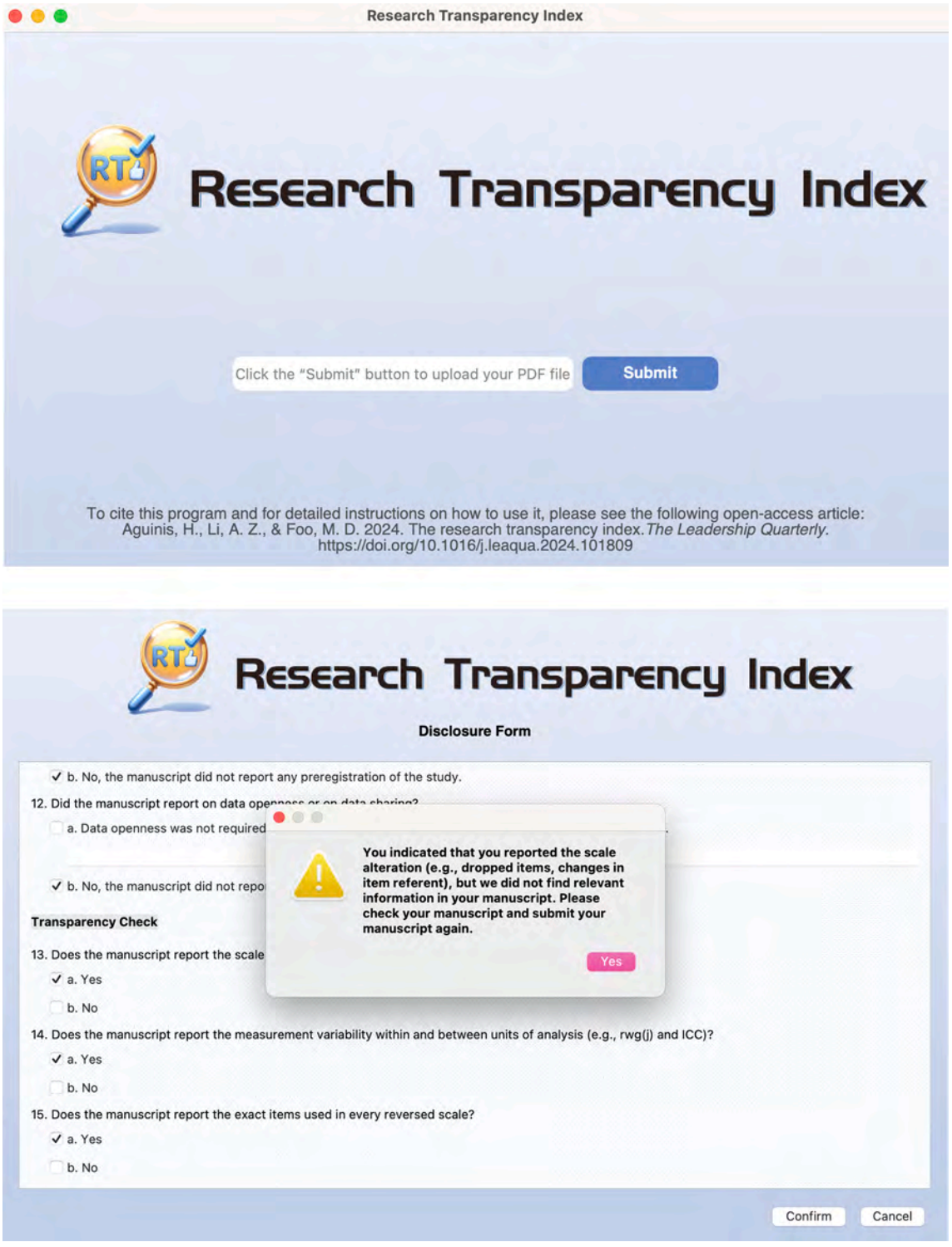

Fig. 1. The Research Transparency Index (RTI) Manuscript Submission and Disclosure Form Interface.

be evaluated solely through software, we recommend including a self-declaration form when utilizing RTI. This practice is advantageous as it allows for assessing transparency issues that may not be readily apparent solely by examining the manuscript, such as excluding unsupported hypotheses from the submitted manuscript. This procedure aligns with established requirements in many journals, where authors are asked to disclose various aspects of their research (e.g., all authors' participation in manuscript production, their awareness of the submission, and the originality of data). For instance, LQ's submission requirement for authors asks, "Do you follow the submission requirements as per the 'Guide to Authors' and, in particular, the 'Data Reporting' guidelines of the journal?" ShinyApps' *Transparency Checklist* also asks that authors confirm the following statement: "The manuscript fully describes how participant dropout was handled (e.g., replaced, omitted, etc.)." The *Transparency Checklist* is a form developed by the ShinyApps platform to help researchers improve and document the transparency of their research. It provides a comprehensive list of items for researchers to self-report, generating a transparency report that can

Table 3
Transparency Disclosure Questions.

Before we proceed with your paper, please respond to the following questions and statements:

General Information

1. By submitting the manuscript to the Research Transparency Index (RTI), I acknowledge that RTI reserves the right to verify the accuracy of the information I have provided. I agree that if requested, I will supply the data to journal staff to verify the results.
   a. Yes
   b. No (If no, please justify)
2. I confirm that I have read and understood the definitions of and requirements concerning plagiarism as defined by *The Leadership Quarterly* author guidelines (https://www.elsevier.com/journals/the-leadership-quarterly/1048-9843?generatepdf=true) and the Elsevier ethical guidelines (https://www.elsevier.com/about/policies/publishing-ethics#Authors), and I acknowledge that I have not committed any form of plagiarism when writing the paper.
   a. Yes
3. I confirm that the research meets the Elsevier ethical guidelines and that I have obtained Institutional Review Board human subjects' approval where appropriate.
   a. Yes
4. I confirm that all my co-authors have verified the data's integrity and the results' accuracy.
   a. Yes
5. If the manuscript is accepted or has been accepted for publication, I agree that this manuscript's RTI score can be displayed publicly.
   a. Yes
   b. No

Paper Transparency

6. Did the manuscript report any post-hoc analyses?
   a. It is not necessary for the manuscript to contain any post-hoc analyses.
   b. The manuscript did not report necessary post-hoc analyses that have been conducted.
7. Did the manuscript report any adapted items?
   a. When collecting data, the original items were used, and none were adapted.
   b. When collecting data, at least one scale was adapted from an original scale, but the manuscript did not report detailed information.
8. Did any of the scales used in the manuscript contain reverse-coded items?
   a. None of the scales used in the manuscript contained reverse-coded items.
   b. At least one of the scales contained reverse-coded items, but the manuscript did not indicate the items.
9. Did any of the scales used in the manuscript contain multiple items?
   a. None of the scales used in the manuscript contained multi-items.
   b. At least one of the scales contained multi-items, but the manuscript did not report any inter-rater agreement-related information.
10. Did the manuscript report any discussion regarding range restriction?
    a. Range restriction analysis was not conducted; therefore, the manuscript does not need to report a discussion regarding range restriction.
    b. No, the manuscript did not report any discussion regarding range restriction.
11. Did the manuscript involve preregistration?
    a. Preregistration was not required; therefore, the manuscript does not need to report any preregistration.
    b. No, the manuscript did not report any preregistration of the study.
12. Did the manuscript report on data openness or data sharing?
    a. Data openness was not required; therefore, the manuscript does not need to report on data.
    b. No, the manuscript did not report on data openness or on sharing data.

be saved and shared publicly. These additional self-declaration questions, which are straightforward to implement (Aczel et al., 2020; Aguinis et al., 2021), can be integrated into existing disclosure forms. The following section of our article provides a detailed explanation of the RTI disclosure form, and the list of questions in the disclosure form is presented in Table 3.

The disclosure (aka self-declaration) form consists of general, paper transparency and self-check questions. In the general questions section, authors are prompted to indicate their awareness of transparency-related information and terms. If authors respond "no" to specific questions, RTI will not generate results for them. An example of a general question that a journal may choose to use is: "By submitting the manuscript to the Research Transparency Index (RTI), I acknowledge that RTI reserves the right to verify the accuracy of the information I have provided. I agree that if requested, I will supply the data to journal staff to verify the analyses and results."

First, authors must declare whether they have written their paper transparently, followed by answering seven paper transparency questions. As discussed earlier, the transparency questions prompt authors to declare and explain the inclusion of important research details, particularly for transparency issues that the software cannot verify. An example of a multiple-choice transparency question is: "Did the manuscript report any adapted items?" The available choices are:

a) Original items were used during data collection, and no adaptation was made.
b) At least one scale was adapted from an original scale during data collection, but the manuscript does not provide detailed information. (If authors choose this option, they must provide appropriate explanations for not including the requested information).

After the general and paper transparency question sections, there are three self-check questions. These self-check questions aim to detect whether the authors have provided honest and accurate responses in the declaration form. An example of a self-check question is: "Have you reported information regarding sensitivity analysis in your manuscript?" In the case where the information provided by authors in the form does not align with the information detected by RTI during the assessment. The tool will identify a potential transparency issue and generate an appropriate prompt, such as: "Your answer does not match the reporting in your manuscript. Please revise your manuscript."

After submitting the pdf file, users can choose the RTI report they wish to view, either for authors and students or for reviewers and editors. A transparency index and report are generated based on their selection. The transparency report consists of three elements. The transparency index for each criterion indicates how well the authors have reported the corresponding research details. For instance, if authors fail to report the exact items used in reversed scales, they will receive 0 points. On the other hand, if authors provide partial

information about the reversed scales (such as reporting only one reversed item when two are included), they will be assigned half of the total points for this criterion. The transparency level for each criterion serves as an additional indicator of transparency. A lower number of points assigned indicates lower transparency. This indicator consists of four levels, ranging from 0, indicating minimal transparency, to 3, which signifies a very high level of transparency.

For criteria that solely require determining their presence (e.g., whether the manuscript has reported the location of data collection), we assign a total transparency score of 2 points. If not detected, the score is 0, if detected, it is 2. For other criteria requiring transparent reporting with explanations or specific descriptions (e.g., whether the manuscript reported both supported and unsupported hypotheses for post-hoc hypotheses with detailed explanations, if any), we assign scores ranging from 0 to 4. If there is no relevant content in the paper, the transparency score for that criterion is 0 points. If the relevant content is detected and necessary details are provided, the score for that criterion is a full 4 points. If only some requirements are met or the criterion name is mentioned without necessary explanations or descriptions, a partial score of 2 points is assigned. We intentionally used 0, 2, and 4 instead of 0, 1, and 2 for each criterion to widen the score differentiation and prepare for finer score adjustments in future tool updates (Aguinis, Pierce, & Culpepper, 2009). The total possible score for all criteria is currently 50 points. A score ranging from 0 to 24 points suggests very low transparency, indicating a critical need for the authors to provide more research details before peer review. A score of 25 to 34 points indicates low transparency, suggesting the manuscript may benefit from additional reporting of research details. Scores between 35 and 44 points represent moderate transparency, indicating that some improvement in reporting research details could be beneficial. Finally, a score of 45 to 50 points signifies high transparency, suggesting that the manuscript is already quite transparent. Table S1 and S2 in the online supplement include a detailed description of the scoring system. In our illustration, the high, medium, and low cutoffs for transparency scores are not the final absolute cutoffs. Journals can customize these initial cutoffs based on their requirements for research transparency and transparency norms (which are likely to evolve over time).

RTI also provides suggestions for improving transparency and lists sample articles with high levels of transparency. These sample articles can serve as references for users to understand how to improve transparency in their manuscripts (the list of sources used as exemplars of high levels of transparency is included in Table S3 in the online supplement). For example, a suggestion regarding reporting reversed scales may be provided:

> "You may want to include information about the reversed items. For instance, if you have used any reversed items, we recommend indicating which items were reversed. For guidance on improving transparency in reporting reversed items, refer to Perkins and Hartless (2002)."

By presenting transparency indexes, reports, and suggestions to users, RTI assists reviewers and editors in automatically identifying missing details, thereby enhancing the efficiency of manuscript evaluation. Importantly, RTI is an *informative and developmental tool that helps authors and students understand the steps they can take to enhance research transparency, thereby increasing the replicability and reproducibility of their manuscripts*.

The design was based on traditional UI design principles (e.g., McKay, 2013), prioritizing a user-centered approach, simplicity, clarity, and consistency. It aims to create visually appealing, accessible, responsive interfaces that seamlessly guide users through an application or website. The primary color scheme draws inspiration from the prevalent colors used in most educational websites and tools, predominantly using clean tones like blue. In addition to using the traditional principles, we also invited researchers to use the tool. They generally considered the tool's design to be user-friendly and straightforward. Also, the RTI platform complies with several key requirements based on Web accessibility (U.S. Department of Justice, 2022). It adheres to the guidelines for text alternatives and captions, primarily relying on text without incorporating videos or images. Additionally, the platform meets the color contrast requirements, ensuring readability for users with visual impairments. Furthermore, considering the tool's potential integration into a resource page with keyboard navigation capabilities, it aligns well with ADA standards for accessibility.

On a related note, AI-related tools, such as ChatGPT, have been extensively developed and utilized for various research purposes. ChatGPT is a general language model that provides conversational responses and generates input-based text. However, our testing with ChatGPT has shown that it does not possess the same functionalities as RTI. While ChatGPT can provide information and answer questions on various topics, it lacks the specific domain expertise and structured evaluation capabilities of RTI (Budhwar et al., 2023). For example, ChatGPT can offer a rough report on research transparency without specific criteria and without a detailed analysis of each dimension. Even when provided with detailed criteria, it often makes mistakes and may provide contradictory answers due to its limited knowledge and incomplete algorithms regarding research transparency. Furthermore, while ChatGPT can sometimes calculate a score based on provided criteria, its reliability is inconsistent and may provide a vague assessment instead. Therefore, despite the value of other AI-related tools, RTI, a tool specifically designed to automatically detect the transparency of research in articles, plays a crucial role in promoting high-quality research, reducing errors, and enhancing transparency.

Using the RTI: Two illustrations

We selected two LQ articles published in 2020 and 2015 for illustrative purposes. The top panel of Fig. 2 includes a screenshot of a portion of the RTI report for authors and students for the 2020 article, indicating a low transparency index (23.08 %). The bottom panel of Fig. 2 shows a screenshot with a portion of the RTI report for reviewers (and editors) for the 2015 article, with a moderate transparency index (57.69 %). In the first article, as shown in the top panel of Fig. 2, the report indicates Level 0 for the theory research stage, Level 2 for the design research stage, Level 1 for the measurement research stage, Level 1 for the data analysis research stage, and Level 1 for the reporting of the results research stage (although the figure shows the Data Analysis and Reporting Style sections of the report only). This report indicates a low level of transparency in the paper. In the second paper, the report in the bottom panel of Fig. 2 indicates Level 0 for the theory research stage, Level 2 for the design research stage, Level 1 for the measurement research stage, Level 2 for the data analysis research stage, and Level 3 for the reporting of results the research stage (although the figure includes screenshots of the last two stages only). From the indicated levels for each research stage, we can see that this paper has a higher overall transparency level than the first paper (i.e., a moderate level). The report provides corresponding suggestions to assist reviewers and editors in providing transparency-related feedback to the authors.

Please note that RTI scores are context-specific and likely to evolve with a field's and journal's norms. For example, if transparency becomes the norm, a very high score (e.g., 90 %) might be considered "'adequate." On the other hand, for a journal and field for which

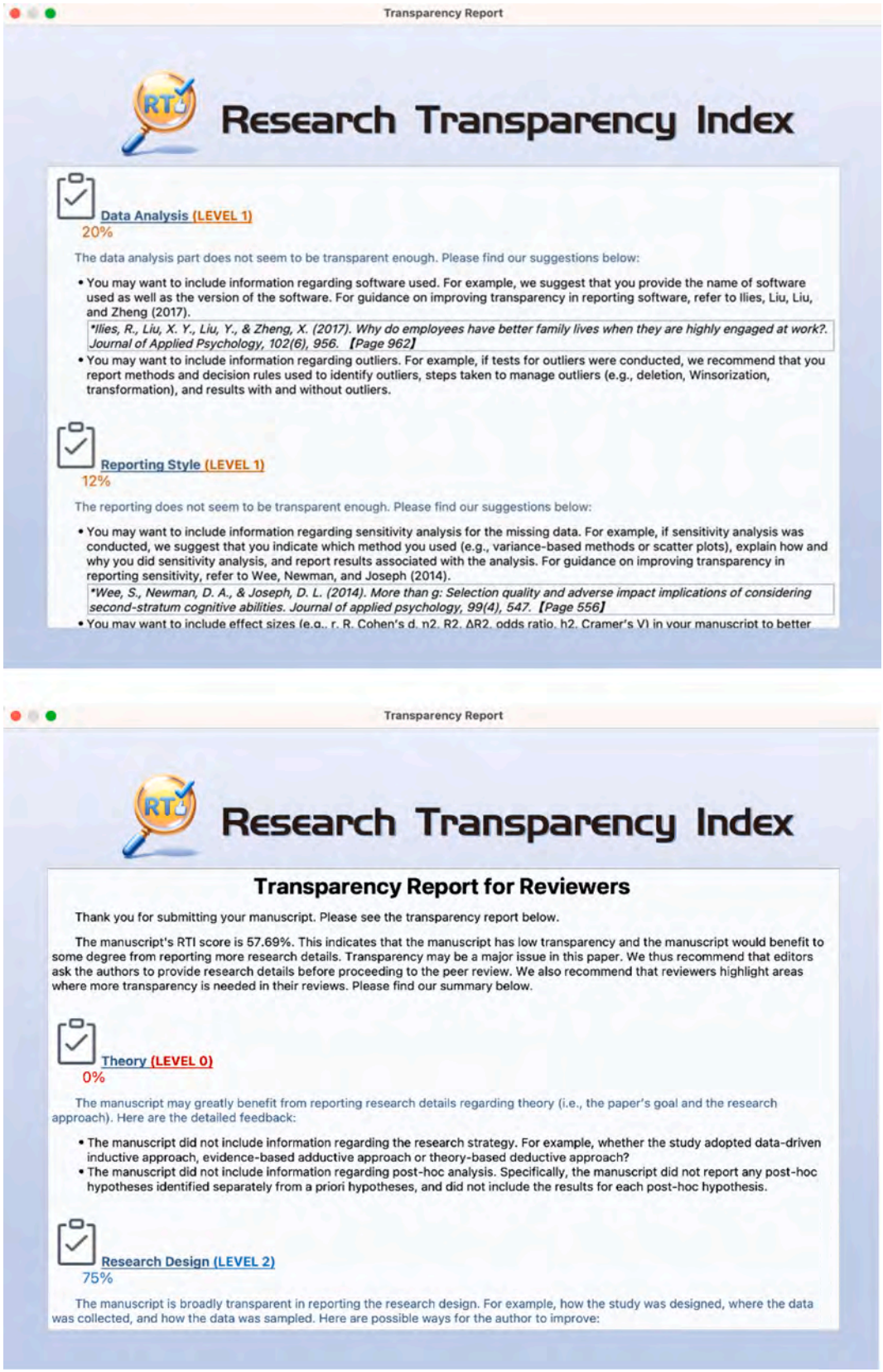

Fig. 2. Examples of RTI-generated Transparency Report for Two *Leadership Quarterly* Illustrative Articles. *Note*. The top panel is an example of Article 1′s transparency report for authors and students with a very low index, and the bottom panel is an example of Article 2′s transparency report for reviewers and editors with a moderate index. The citations for these illustrative LQ articles published in 2020 and 2015 are available from the authors upon request.

transparency is not firmly established as a necessity, a score of, for example, 60 % might be considered adequate. For these reasons, RTI provides a score, but the evaluative judgment of whether a particular one qualifies for one threshold or another is journal, field, and time-dependent.

Validity Evidence for RTI

To gather concurrent validity evidence about the RTI, we conducted two studies to compare the results of manual checking with those obtained using RTI. In each validity study, we recruited three management Ph.D. candidates, totaling six participants, to evaluate the transparency ratings of four selected papers based on our criteria. We then calculated $r_{wg(j)}$ to assess the internal consistency between the ratings the three Ph.D. candidates provided and those generated by RTI for each study.

We provided all six Ph.D. candidates with instructions on our research focus, which centered on analyzing published journal articles. We informed them that our objective was to conduct a peer review of papers to assess the extent to which each paper incorporated specific criteria. Initially, we requested the candidates to provide personal information and indicate their relevant review-related experiences. This information included their highest educational degree, whether they had presented a peer-reviewed paper, published an article in a peer-reviewed journal, and served as a reviewer for an academic conference or a peer-reviewed journal. All six Ph.D. candidates had completed at least two courses in quantitative research methodology and received instruction on transparency through their coursework or by reading relevant articles. We confirmed their understanding of the definition of transparency (i.e., "the degree of detail and disclosure about the specific steps, decisions, and judgment calls made during a scientific study"; Aguinis et al., 2018, p. 83), the criteria involved to evaluate transparency, and how to identify them. Four of the candidates had previous experience as reviewers for academic conferences, four had presented papers at conferences, and two had published at least one article in peer-reviewed journals. The participants were instructed to read the criteria descriptions and complete a criterion rating form. To facilitate their understanding of the rating mechanism, we provided an example and emphasized that there were no right or wrong answers.

Validity study #1

In the first validity study, we recruited three management Ph.D. candidates from universities in Singapore as participants, and each of them was assigned four articles. We initially selected 30 published articles from 2000 to 2020, as transparency issues have gained increasing attention over the past two decades. We excluded articles published more than 20 years earlier to maximize currency and relevance, as they might have exhibited very low transparency. The four articles assigned to the reviewers were randomly chosen from *Journal of Applied Psychology* (2019), *Journal of Business Venturing* (2020), *Journal of International Business Studies* (2018), and *Journal of Management* (2010). We purposely chose these journals across fields to assess RTI's ability to assess transparency across management domains and adjacent fields. Table 4 shows the descriptive statistics for the two validity studies.

The participants were first instructed to score 17 transparency criteria using a scale ranging from 1 (criterion not present) to 5 (criterion certainly present), with an option for "NA" (not applicable) if the criterion did not apply to a specific paper. To simplify the initial version of RTI and reduce the reviewers' time and effort, we selected three to four criteria from each previously discussed research stage: theory, research design, measurement, data analysis, and reporting style. For this first validity study, the focus was on the following criteria: research strategy, post hoc analyses, type of research design, type of data collection, country of data collection, type of sampling method, scale alteration, reversed item, measurement variability, range restriction, analytic method, name and version of software, outlier detection, sensitivity analysis, response rate, reporting *p*-values, and reporting coefficients. Table 5 presents the 17 criteria included in the first validity study and their descriptions.

We calculated $r_{wg(j)}$ to assess the internal consistency between the raters and RTI (Table S4 in the online supplement includes the ratings). The $r_{wg(j)}$ for Article 1 was 0.88, with "NA" replaced with 1. For Article 2, $r_{wg(j)}$ was 0.84, and "NA" was replaced with 1. The $r_{wg(j)}$ for Article 3 was 0.65, and for Article 4, it was 0.78. In these calculations, 1 represented "criterion not present," while NA denoted "not applicable," indicating the absence of relevant information in the paper.[2] The average $r_{wg(j)}$ was 0.79 ($r_{wg(j)\min} = 0$, $r_{wg(j)\max} = 1$, $r_{wg(j)mean} = .52$, $r_{wg(j)SD} = .46$), considered acceptable (Lance et a., 2006).[3] In our study, two primary scenarios contribute to an $r_{wg(j)}$ of 0. First, discrepancies in understanding or personal biases may affect a rater's scoring. The second scenario involves raters not thoroughly reviewing materials, resulting in ratings that starkly contrast with those of their peers. These situations further underscore the importance of the RTI tool, which, when utilized adequately with well-defined criteria and algorithms, can significantly enhance research transparency. Furthermore, the current RTI framework allows only limited scoring outcomes (yes, no, partial), preventing the assignment of intermediate scores like 2 or 4, thus impacting $r_{wg(j)}$ values. This limitation suggests that $r_{wg(j)}$ scores might not entirely reflect the tool's validity concerns. With the integration of more sophisticated AI technologies in the future, the RTI is expected to offer more detailed ratings, thus providing more precise insights into the research transparency and enhancing the interpretive value of $r_{wg(j)}$ scores.

It is important to distinguish between the scale points used in the validity studies and the levels in the RTI scoring system, as they represent two different concepts. A traditional 5-point scale is utilized for the validity test, which is familiar to reviewers for easy scoring. On the other hand, the levels in the RTI scoring system signify the transparency level of criteria. Four levels have been established because each criterion's score is an even number, as explained in the section "Using the RTI" regarding the scoring system, hence the use of even-numbered levels.

Validity study #2

Based on the results from the first validity study, we made further revisions to improve RTI's reliability. In addition, we enhanced the clarity of criterion descriptions in the second validity study to minimize ambiguity. We incorporated prefix terms like "explicitly state" to achieve this aim and provided more detailed explanations for certain criteria. Furthermore, we introduced one additional critical criterion, effect size, in the second validity study. Effect sizes complement statistical significance testing, and many journals require their reporting in manuscripts. Table 6 contains the descriptions of the 18 criteria assessed in the second validity study. As in the first study, we invited three additional management Ph.D. candidates to review four articles. The selected articles were from the following journals: *Journal of Management* (2020), *Journal of International Business Studies* (2015), *Journal of Business Venturing* (2011), and *Academy of Management Journal* (2009). Similar to Validity Study #1, we purposely chose these journals across

[2] We included the "not applicable" (NA) option because, for some papers, reviewers may not be sure whether to assign the lowest score if a paper does not report relevant information. For instance, in the case of altered items, some manuscripts may mention using a well-developed scale but fail to specify whether any items were altered. In such cases, reviewers who are uncertain about the scoring can choose NA. Therefore, in the actual calculation of $r_{wg(j)}$, we substituted ''NA'' with the lowest score, "1", as NA generally indicates that the paper does not mention relevant information.

[3] We conducted the $r_{wg(j)}$ analyses assuming a triangular distribution. The results for both studies yielded values of 1, indicating that our $r_{wg(j)}$s, calculated with the assumption of a normal distribution, are not elevated.

Table 4
Descriptive Statistics for Two Validity Studies.

	Validity Study #1								Validity Study #2							
	Article 1		Article 2		Article 3		Article 4		Article 1		Article 2		Article 3		Article 4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Research strategy	3.00	2.31	2.75	2.06	2.50	1.91	2.25	1.89	1.50	1.00	1.50	1.00	1.50	1.00	4.50	1.00
Post hoc analysis	4.00	1.15	1.25	0.50	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.50	1.00
Type of research design	5.00	0.00	4.00	2.00	3.50	1.91	3.75	1.89	3.75	0.96	3.00	1.63	3.50	1.29	5.00	0.00
Type of data collection	4.75	0.50	5.00	0.00	4.50	1.00	4.75	0.50	4.25	0.96	3.75	1.26	4.75	0.50	5.00	0.00
Country of data collection	5.00	0.00	5.00	0.00	5.00	0.00	5.00	0.00	5.00	0.00	5.00	0.00	5.00	0.00	4.75	0.50
Type of sampling method	2.50	1.91	4.00	2.00	3.00	1.83	3.75	1.89	2.00	0.82	2.25	1.50	2.00	1.15	3.50	1.91
Scale alteration	2.00	2.00	4.50	1.00	3.25	1.71	1.00	0.00	3.25	2.06	2.00	2.00	3.25	1.71	2.00	2.00
Reversed item	1.00	0.00	2.25	1.89	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	3.75	1.89
Measurement variability	1.00	0.00	1.75	1.50	1.75	1.50	1.00	0.00	1.00	0.00	1.25	0.50	1.75	1.50	2.75	2.06
Range restriction	1.00	0.00	1.00	0.00	1.00	0.00	1.50	1.00	3.00	2.31	1.00	0.00	1.00	0.00	2.00	2.00
Analytic method	4.75	0.50	3.75	0.96	4.50	1.00	3.00	2.31	4.25	0.50	3.75	0.96	5.00	0.00	5.00	0.00
Name of software and version of software	1.00	0.00	2.25	1.89	4.00	1.15	3.50	1.73	5.00	0.00	2.00	2.00	3.50	1.91	4.00	2.00
Outlier detection	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	3.00	2.31
Sensitivity analysis	1.50	1.00	2.50	1.91	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	2.25	1.89
Response rate	3.00	2.31	5.00	0.00	5.00	0.00	1.00	0.00	5.00	0.00	1.25	0.50	5.00	0.00	2.00	2.00
Effect Size	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a	5.00	0.00	5.00	0.00	3.00	2.31	4.00	2.00
Reporting *p*-values	4.50	1.00	3.00	2.31	3.50	1.91	3.00	1.63	3.50	1.00	2.75	1.26	3.75	1.89	5.00	0.00
Reporting coefficients	3.00	2.31	4.00	2.00	4.00	2.00	4.50	1.00	4.00	2.00	4.00	2.00	3.50	1.91	4.25	0.96

Note. Validity Study #1: Article 1: *Journal of Applied Psychology* (2019), Article 2: *Journal of Business Venturing* (2020), Article 3: *Journal of International Business Studies* (2018), Article 4: *Journal of Management* (2010). Validity Study #2: Article 1: *Journal of Management* (2020), Article 2: *Journal of International Business Studies* (2015), Article 3: *Journal of Business Venturing* (2011), Article 4: *Academy of Management Journal* (2009). The citations for the articles used in the validity studies are available from the authors upon request. n.a.: Not applicable because Validity Study #1 did not include assessing effect size transparency.

fields to assess RTI's ability to assess transparency across management domains and adjacent fields.

After the modifications made to the RTI and the revision of criterion descriptions, the $r_{wg(j)}$ for Article 1 was 0.83, with "NA" replaced with 1 (Table S5 in the online supplement includes the ratings). For Article 2, $r_{wg(j)}$ was 0.95, and "NA" was replaced with 1. The $r_{wg(j)}$ for Article 3 was 0.92, and for Article 4, it was 0.88. In these calculations, 1 represented "criterion not present," while NA denoted "not applicable," indicating the absence of relevant information in the paper. The average $r_{wg(j)}$ was 0.90 ($r_{wg(j)\min} = 0$, $r_{wg(j)\max} = 1$, $r_{wg(j)mean} = .54$, $r_{wg(j)SD} = .44$), considered acceptable according to Lance et al. (2006). This average $r_{wg(j)}$ score was higher than that of validity study #1, indicating improved reliability after the modifications and revisions were implemented.

One of us reviewed the ratings and compared the results between the RTI report and our manual review. We discovered that for criteria requiring specific numerical information (such as *p*-values, information about the software used, and reporting coefficients), RTI achieved more accurate results than manual ratings. During the manual review, the

Table 5
Transparency Criteria Descriptions for Validity Study #1.

Criterion	Description
Theory	
1. Research strategy	Did the article specify the research strategy (e.g., inductive, deductive, abductive)?
2. Post hoc analysis	Did the article identify and report any post-hoc hypotheses separately from a priori hypotheses? If yes, did the article report both supported and unsupported hypotheses?
Research Design	
3. Type of research design	Did the article include the type of research design (e.g., passive observation, experimental)?
4. Type of data collection	Did the article report the form of data collection procedure (e.g., surveys, interviews)?
5. Country of data collection	Did the article report the location of data collection (e.g., North America/China)?
6. Type of sampling method	Did the article report the type of sampling method (e.g., purposeful, snowball, convenience)?
Measurement	
7. Scale alteration	**If scales were altered,** did the article report how and why (e.g., dropped items, changes in item referent)? Did the article provide psychometric evidence regarding the altered scales (e.g., feature-related validity)?
8. Reversed item	Did the article report the exact items used in the reversed scale, **if any?**
9. Measurement variability	**If scores are aggregated**, did the article report measurement variability to justify aggregation (e.g., $r_{wg(j)}$, ICC)?
10. Range restriction	**If range restriction was assessed,** did the article specify the type of range restriction and provide a rationale for the decision to correct or not correct? If corrected, did the article identify how (e.g., type of correction used, sequence, and formulas) and report observed effects and those corrected for range restriction?
Information on Data Analysis	
11. Analytic method	Did the article report the specific analytical method used and why it was chosen (e.g., EFA versus CFA; repeated measures ANOVA using conventional univariate tests of significance, structural equation modeling, etc.)?
12. Name of software and version of software	Did the article report the software used, including which version and the name of the software?
13. Outlier detection	If tests for outliers were conducted, did the article report methods and decision rules used to identify outliers, steps (if any) taken to manage outliers (e.g., deletion, Winsorization, transformation), the rationale for those steps, and results with and without outliers?
Information on Reporting	
14. Sensitivity analysis	Did the article report results of missing data analysis (e.g., sensitivity analysis), if any? Did the report provide details regarding the sensitivity analysis?
15. Response rate	Did the article report the exact response rate?
16. Reporting *p*-values	Did the article report exact *p*-values to two decimal places?
17. Reporting coefficients	Did the article report and clearly identify coefficients as either unstandardized or standardized?

Table 6
Transparency Criteria Descriptions for Validity Study #2.

Criterion	Description
Theory	
1. Research strategy	Did the article specify the research strategy (e.g., explicitly indicate inductive, deductive, or abductive research)?
2. Post hoc analysis	Did the article identify and report any post-hoc hypotheses separately from a priori hypotheses**?** If yes, did the article report both supported and unsupported hypotheses? *(Highest Possible Score: post-hoc analysis reported with detailed explanation.)*
Research Design	
3. Type of research design	Did the article include the type of research design (e.g., explicitly indicate passive observation, experimental)?
4. Type of data collection	Did the article report the form of data collection procedure (e.g., explicitly state survey interviews)?
5. Country of data collection	Did the article report the location of data collection (e.g., North America/China)?
6. Type of sampling method	Did the article report the type of sampling method (e.g., explicitly state purposeful, snowball, convenience)?
Measurement	
7. Scale alteration	**If scales were altered,** did the article report how and why (e.g., dropped items, changes in item referent)? Did the article provide psychometric evidence regarding the altered scales (e.g., feature-related validity)? *(Highest Possible Score: altered items and explanations were reported.)*
8. Reversed item	Did the article report reverse scales, **if any?**
9. Measurement variability	**If scores are aggregated**, did the article report measurement variability to justify aggregation (e.g., $r_{wg(j)}$, ICC)?
10. Range restriction	**If range restriction was assessed,** did the article specify the type of range restriction and provide a rationale for the decision to correct or not? If corrected, did the article identify how (e.g., type of correction used, sequence, and formulas) and report observed effects and those corrected for range restriction? *(Highest Possible Score: range restriction analysis reported with detailed explanation.)*
Information on Data Analysis	
11. Analytic method	Did the article report specific analytical methods (e.g., EFA versus CFA; repeated measures ANOVA using conventional univariate significance tests, structural equation modeling, etc.)?
12. Name of software and version of software	Did the article report the software used, including which version and the name of the software? *(Highest Possible Score: both the name of the software and the version of the software were reported.)*
13. Outlier detection	If tests for outliers were conducted, did the article report methods and decision rules used to identify outliers, steps (if any) taken to manage outliers (e.g., deletion, Winsorization, transformation), the rationale for those steps, and results with and without outliers? *(Highest Possible Score: outlier analysis reported with detailed explanation.)*
Information on Reporting	
14. Sensitivity analysis	Did the article report results of missing data analysis (e.g., sensitivity analysis), if any? Did the report include details regarding the sensitivity analysis? *(Highest Possible Score: A sensitivity analysis was reported with a detailed explanation.)*
15. Response rate	Did the article report the exact response rate?
16. Effect size	Did the article report any effect sizes, such as confidence intervals, Cohen's *d*, *r*, and R^2?
17. Reporting *p*-values	Did the article report exact *p*-values to two decimal places *(Highest Possible Score: all p-values were reported in the exact format, and authors should not report p-values compared to cut-offs* (e.g., $p < .05$ *or* $p < .01$*)?)*
18. Reporting coefficients	Did the article report and clearly identify coefficients as either unstandardized or standardized?

Note. As explained, we added a new criterion of effect size (#16) in the second validity study because effect sizes complement statistical significance testing, and many journals require their reporting in manuscripts.

raters tended to assign higher scores for *p*-value-related information, even if the *p*-values were not explicitly stated (e.g., "$p < 0.5$"). In contrast, RTI followed strict criteria and conducted a comprehensive check of all *p*-values in the submitted papers, leading to more accurate assessments. Thus, RTI demonstrated greater accuracy than manual reviewers regarding numbers and specific information-focused criteria.

Furthermore, during the two validity studies, we observed that some reviewers expressed uncertainty when assigning scores to certain criteria. For example, regarding the criterion of reversed items, several reviewers were unsure whether to give a score of 1 or mark "NA" because they were uncertain whether including reversed items in the original study was required. For example, if the scale does not include reversed items, it remains unclear to reviewers whether it is still necessary to address this issue. This observation emphasizes the importance of authors providing transparency declarations. Having clear information from the authors can help mitigate uncertainties and improve the accuracy of the transparency assessment.

In RTI v. 1.0, we selected 18 out of 31 criteria based on LQ editorials and articles from the editorial team of George C. Banks (e.g., Antonakis, 2017; Wulff et al., 2023) as well as the checklist for primary quantitative studies provided by an open-science tool like ShinyApps suggested by *The Leadership Quarterly* (Aczel et al., 2020). These 18 criteria, italicized in Table 1, represent the most frequently discussed issues and are typically expected to be included in manuscripts that use quantitative methods. For instance, under the reporting style research stage, we included criteria such as missing data analysis, effect sizes, *p*-values, and whether reported coefficients are indicated as standardized or unstandardized. The remaining 13 criteria are suggested for authors to self-check using the declaration form. However, they can be incorporated into future versions of RTI.

RTI serves as a valuable tool for detecting research transparency, and we encourage researchers to expand its application to other types of papers, including those with multiple studies, experimental studies, qualitative studies, and meta-analyses. Each type of study will require different transparency criteria. For example, transparency in meta-analyses (Wulff et al., 2023) and qualitative research (Knight et al., 2022) will involve distinct criteria. As new research methods emerge, such as AI and machine language, the transparency criteria may require revision. In such cases, the RTI tool can be updated accordingly, and those updates will be described at https://www.hermanaguinis.com. Moreover, we also make the Python code available so others can implement updates as well.

Conclusions

Several journals and professional organizations are implementing strategies to enhance research transparency. For example, *The Leadership Quarterly* embraces open science practices, such as pre-registration and using ShinyApps' checklists (Aczel et al., 2020), requiring authors to provide extensive information during the paper submission process, including all control variables in the correlation table. *Journal of Management* has revised the reviewer evaluation form and expanded and clarified data disclosure requirements for all methodologies, ensuring that authors provide extensive information during submission (Bergh & Oswald, 2020). *Academy of Management Journal* encourages pre-registration, data-sharing, and the use of reporting guidelines (DeCelles et al., 2021). *Journal of Applied Psychology* has developed a reporting checklist to guide authors in improving the transparency of their studies. As journals demand more detailed transparency requirements, authors and reviewers may face increased burdens in ensuring data transparency. RTI can help alleviate this burden by automatically checking for key inputs critical for transparency.

We emphasize that compliance with the specific requirements of each journal is essential. While RTI allows journals to establish their policies regarding the tool's use (e.g., mandatory, voluntary, or for developmental purposes before submission) and decide which criteria to include in evaluating transparency, RTI can only serve as an exclusively developmental tool. The assessment of transparency levels is not an absolute endeavor. Instead, it is contingent upon the specific journal and the prevailing norms within the respective academic field. Thus, while RTI provides a quantitative score to gauge transparency, *the evaluative judgment of whether a particular score qualifies as meeting a specific threshold or another is contingent on the unique characteristics of the journal, the dynamics of the academic field, and the prevailing norms at a given point in time*. Moreover, *RTI should not be the exclusive arbiter of a manuscript's transparency level and cannot replace human judgment when examining it*. Therefore, authors should carefully review and adhere to the author guidelines provided by each journal to ensure alignment with publication standards.

RTI is designed to benefit various user groups, including authors, students, reviewers, and editors. Authors can utilize RTI to assess the adequacy of their information and use the tool's suggestions to improve the transparency of specific areas, such as measurement and reporting styles. In an educational context, students can utilize RTI as a self-learning tool to promote transparent research outcomes. In addition, instructors and thesis and dissertation committee members can use RTI to guide students on the principles and elements of transparency. Additionally, RTI can assist reviewers and editors in streamlining the review process by providing an automated assessment of a manuscript's transparency. Editors can use the RTI score and report to determine whether a submitted paper should proceed to the review stage or be returned to the authors for revision without manual verification of including all required information. Finally, organization and policy decision-makers will benefit from using RTI to gather information on the relative transparency of research that may serve as input for creating policies and making decisions. Even with little knowledge in a specific field, managers and policy-makers can utilize RTI to understand whether specific research studies they may want to rely on are transparent.

In addition, the RTI can be used to research the domain of transparency. For example, it can answer questions such as: Which journals are more transparent and regarding which specific criteria? Which domains are more transparent (e.g., organizational behavior vs. human resource management vs. leadership)? Has the degree of transparency changed faster in some journals and domains compared to others? What are the predictors of transparency, and what are the outcomes? For example, are journals with higher transparency reporting more or fewer retracted articles? Have journals that published articles with higher transparency increased their impact factor over time? Overall, we see great potential in using the RTI not only as a developmental tool but also as a tool to collect data on transparency that will allow us to improve our understanding of its antecedents and outcomes at various levels of analysis (e.g., individual research, journal, domain, and field).

In conclusion, developing the RTI tool is valuable to the ongoing efforts to enhance research transparency. By providing an automated assessment of transparency criteria, RTI saves authors, students, reviewers, and editors time while promoting transparent research practices. As journals prioritize transparency, RTI is a powerful resource to support authors in meeting journal requirements and encourages the advancement of rigorous and transparent research. Ultimately, through its user-friendly interface and comprehensive evaluation, RTI facilitates a culture of transparency, fostering trust and credibility in the scholarly community and among users of the knowledge we produce (e.g., organization and policy decision-makers).

CRediT authorship contribution statement

Herman Aguinis: Writing – original draft, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Zhuyi Angelina Li:** Writing – review & editing, Writing – original draft, Validation, Software, Formal analysis, Data curation. **Maw Der Foo:** Writing – review & editing, Writing – original draft, Validation, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.leaqua.2024.101809.

References

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., & Wagenmakers, E. J. (2020). A consensus-based transparency checklist. *Nature Human Behaviour, 4*(1), 4–6.

Aguinis, H. (2025). *Research methodology: Best practices for rigorous, credible, and impactful research*. Sage.

Aguinis, H., & Solarino, A. M. (2019). Transparency and replicability in qualitative research: The case of interviews with elite informants. *Strategic Management Journal, 40*(8), 1291–1315.

Aguinis, H., Cascio, W. F., & Ramani, R. S. (2017). Science's reproducibility and replicability crisis: Management is not immune. *Journal of International Business Studies, 48*(6), 653–663.

Aguinis, H., Hill, N. S., & Bailey, J. R. (2021). Best practices in data collection and preparation: Recommendations for reviewers, editors, and authors. *Organizational Research Methods, 24*(4), 678–693.

Aguinis, H., Pierce, C. A., & Culpepper, S. A. (2009). Scale coarseness as a methodological artifact: Correcting correlation coefficients attenuated from using coarse scales. *Organizational Research Methods, 12*(4), 623–652.

Aguinis, H., Ramani, R. S., & Alabduljader, N. (2018). What you see is what you get? Enhancing methodological transparency in management research. *Academy of Management Annals, 12*(1), 83–110.

Aguinis, H., Ramani, R. S., & Cascio, W. F. (2020). Methodological practices in management research: An after-action review of challenges and solutions. *Journal of International Business Studies, 51*(9), 1593–1608.

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5–21.

Banks, G. C., O'Boyle, E., Pollack, J. M., White, C. D., Batchelor, J. H., Whelpley, C. E., Abston, K. A., Bennett, A. A., & Adkins, C. L. (2016). Questions about questionable research practices in the field of management A guest commentary. *Journal of Management, 42*(1), 5–20.

Bergh, D. D., & Oswald, F. L. (2020). Fostering robust, reliable, and replicable research at the Journal of Management. *Journal of Management, 46*(7), 1302–1306.

Bettis, R. A. (2012). The search for asterisks: Compromised statistical tests and flawed theories. *Strategic Management Journal, 33*(1), 108–113.

Budhwar, P., Chowdhury, S., Wood, G., Aguinis, A., Bamberb, G. J., Beltran, J. R., Boselied, P., Cookee, F. L., Decker, S., DeNisi, A., Dey, P. K., Guest, D., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., Rogelberg, S., Saunders, M. N. K., Tung, R. L., & Varma, A. (2023). HRM in the age of generative AI: Perspectives and research direction on ChatGPT. *Human Resource Management Journal, 33*(3), 606–659.

DeCelles, K. A., Howard-Grenville, J., & Tihanyi, L. (2021). From the editors—Improving the transparency of empirical research published in AMJ. *Academy of Management Journal, 64*(4), 1009–1015.

Friedrich, T. L., Byrne, C. L., & Mumford, M. D. (2009). Methodological and theoretical considerations in survey research. *The Leadership Quarterly, 20*(2), 57–60.

Gelles, D. (2023). *When professors cheat: A Harvard star is accused and a culture is questioned*. The New York Times.

Hill, N. S., Aguinis, H., Drewry, J., Patnaik, S., & Griffin, J. J. (2022). Using macro archival databases to expand theory in micro research. *Journal of International Business Studies, 59*(3), 627–659.

Knight, G., Chidlow, A., & Minbaeva, D. (2022). Methodological fit for empirical research in management: A contingency framework. *Journal of International Business Studies, 53*(1), 39–52.

Kraimer, M. L., Martin, X., Schulze, W., & Seibert, S. E. (2023). What does it mean to test theory? *Journal of Management Scientific Reports, 1*(1), 8–17.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods, 9*(2), 202–220.

McKay, E. N. (2013). *UI is communication: How to design intuitive, user-centered interfaces by focusing on effective communication*. Newnes.

O'Boyle, E. H., Jr., Banks, G. C., & Gonzalez-Mule, E. (2017). The chrysalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management, 43*(2), 376–399.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.

Toth, A. A., Banks, G. C., Mellor, D., O'Boyle, E. H., Dickson, A., Davis, D. J., DeHaven, A., Bochantin, J., & Borns, J. (2021). Study preregistration: An evaluation of a method for transparent reporting. *Journal of Business and Psychology, 36*(4), 553–571.

U.S. Department of Justice, Civil Rights Commission (2022, March 18). *Guidance on Web accessibility and the ADA*. https://www.ada.gov/resources/web-guidance/.

Wulff, J. N., Sajons, G. B., Pogrebna, G., Lonati, S., Bastardoz, N., Banks, G. C., & Antonakis, J. (2023). Common methodological mistakes. *The Leadership Quarterly, 34*(1), Article 101677.