

PERSPECTIVE

Are media reports of published research an accurate representation of the research?

Jingqi Yu¹ , Catherine Yeung²  and Dilip Soman¹

¹Rotman School of Management, University of Toronto, Toronto, ON, Canada and ²Department of Marketing, CUHK Business School, Chinese University of Hong Kong, Hong Kong

Corresponding author: Catherine Yeung; Email: cyeung@cuhk.edu.hk

(Received 4 July 2024; revised 5 March 2025; accepted 10 March 2025)

Abstract

It is perhaps uncontroversial to claim that behavioral science research is playing an increasingly important role in practice. However, practitioners largely rely on media reports rather than original research articles to learn about the science. Do these media reports contain all the information needed to understand the nuances of the research? To assess this question, we develop a set of rubrics to evaluate the fidelity of the media report to the original research article. As an illustration, we apply these rubrics to a sample of media reports based on several research articles published in one journal and identify common patterns, trends, and pitfalls in media presentations. We find preliminary evidence of low fidelity in presenting participant characteristics, contextual elements, and limitations of the original research. The media also appear to misreport correlational evidence as causal and sometimes miss acknowledging the hypothetical nature of evidence when hypothetical scenarios were used as the sole basis of conclusions. Furthermore, the media often present broad conclusions and personal opinions as directly backed by scientific evidence. To support more discerning consumption of behavioral insights from media sources, we propose a checklist to guide practitioners in evaluating and using information from media sources.

Keywords: behavioral interventions; behavioral science; knowledge translation; media reports; reporting fidelity

Academic research in the behavioral sciences has taken on an increasingly important role in practice in recent years (Thaler and Sunstein, 2021). Behavioral units in government agencies, for-profit corporations, and non-profit organizations are using insights from the behavioral sciences to develop behavior change interventions that have benefited millions of individuals worldwide and resulted in successful business outcomes (Hubble and Varazzani, 2023). The growing adoption of behavioral insights has also meant that practitioners need to stay informed about the latest research. However, the path from research to application is not as straightforward as one might assume. A recent survey showed that only 41% of practitioners read original research articles,

while the remaining majority rely on media reports, such as social media, popular press articles, and podcasts (collectively referred to as media reports; 49%), or other non-academic sources (10%) to learn about behavioral science research (Yu and Feng, 2024). This heavy reliance on media for consuming scientific findings raises questions about the validity of the knowledge transfer process and the fidelity of the reporting. If interventions are going to be science-based, it is imperative that designers of those interventions get the nuances of the science right!

In this article, we aim to accomplish four main objectives. First, we make the case for why we – as behavioral public policy scientists – need to study how our work gets translated by the media and consumed by practitioners. Second, we introduce a set of rubrics grounded in theory and practice to evaluate how the media presents behavioral research findings. Third, we illustrate the use of the rubrics by conducting a thorough evaluation of a small sample of research articles published in the journal *Natural Human Behaviour* (NHB) to provide some preliminary evidence about media representations of scientific knowledge. Finally, we propose a checklist that can serve as a decision-support tool to guide practitioners in their consumption of knowledge from media sources.

Fidelity of media reporting – why does it matter?

The role of media reporting in the political economy of science

Research findings in behavioral science are part of a larger system that encompasses the political economy of science. This system involves the production of knowledge based on scientific research, the transmission of such knowledge to potential stakeholders – policymakers, practitioners, and members of the general public as opinion influencers and voters – and the use of such findings as evidence to support policymaking (Crowley *et al.*, 2021; Hjort *et al.*, 2021; Briscese and List, 2024; Toma and Bell, 2024; Stefano *et al.*, 2024). In this system, researchers conduct studies, and media reporters communicate the findings to the public and practitioners, shaping beliefs, trust, and support of the general public, and influencing policy decisions (Briscese and List, 2024; Toma and Bell, 2024). Media reports serve as a crucial link, providing quick access to the latest scientific developments for those who may not engage with original research articles. Therefore, research evidence and the parties involved in its dissemination play not only an isolated role in evaluating how an intervention works in a given setting but also a supportive role in helping both policymakers and the public make informed decisions on how interventions can be adopted, maintained, and scaled up (Stefano *et al.*, 2024).

The media that reports scientific findings shapes the public's understanding of the relevance and limitations of scientific findings. When media reports omit crucial details and present research findings as universally applicable, practitioners may be misled into applying interventions in contexts where they are less likely to be effective (List, 2022). Such misinformation not only leads to costly failures but also erodes both the practitioner's and the public's trust in behavioral science, reducing the motivation to implement evidence-based strategies in the long run (Mažar and Soman, 2022). Moreover, inaccuracies in media reports can lead to persistent misconceptions and biased beliefs among the public, which could, in turn, influence public acceptance of policies even though they may be beneficial to society (Stefano *et al.*, 2024). The

consequences of low-fidelity media reporting can be far-reaching, as public opinion and trust in science play a crucial role in the political economy of science.

Contextual sensitivity of behavioral science research

As we pointed out earlier, media reports provide practitioners with quick access to the latest scientific developments, but there is a risk of practitioners misapplying these findings due to a simplistic understanding of the research. This would not be a problem in an ideal world in which the media report comprehensively describes the nuances of the original research, equipping practitioners to judge the applicability of findings to their contexts. In reality, however, media reports are designed to captivate the audience's attention by presenting newsworthy content. They often simplify complex research, focusing on large and important questions, and offer simple and quick 'takeaway' solutions (Kuehn and Lingwall, 2016). In this simplification process, the nuances of the original research may unintentionally be lost.

The effectiveness of behavioral interventions is particularly susceptible to contextual nuances. Recent work has shown that changes in context and target population can render previously successful interventions ineffective elsewhere (List, 2020; Yang *et al.*, 2023). Consider the use of frameworks and guiding principles that distill multiple ideas to help practitioners design interventions; for example, the EAST framework proposed by the Behavioral Insights Team (Behavioral Insights Team, 2014). This framework suggests that to promote positive behavioral change, we should make the desired behavior Easy, make the anticipated reward Attractive, harness the power of Social influence, and nudge at the right Timing when people are receptive (i.e., be Timely). Each of these principles holds up well in some demonstrations, but specific interventions may work or not work, depending on what context it is translated to.

For example, while simplifying information can enhance engagement by reducing the cost of information processing, Shah *et al.* (2024) found that simplifying pension statements could backfire and lead to lower voluntary retirement saving contributions when the simplification drew attention to the fund's low performance. Notably, people did not switch to better-performing investment agents (in that case, the simplification would have still improved the financial return of the individual investors, despite reducing revenue for the underperforming agents) because of the transaction costs involved; instead, the individual investors disengaged and reduced their total contributions. Here, the contextual factors of low fund performance and the transaction costs of switching conspired to create a context in which the 'Easy' principle backfired.

Similarly, increasing the attractiveness of the anticipated reward appears to be a bulletproof strategy to increase the behavior. However, the implementation of this attractiveness principle is highly sensitive to context and population differences in what is considered an 'attractive' benefit. Shah *et al.* (2023) found that framing savings as a way to secure one's family's future increased retirement contributions, but this intervention backfired for younger people who did not have a family (e.g., those under the age of 28).

The well-known social-norming technique, widely adopted by *Opower* in the United States to reduce household energy consumption, was cost-effective despite a voltage drop when scaled up from proof-of-concept studies to the community level (see

Schultz *et al.*, 2007; Allcott, 2011). However, in Germany, where baseline energy consumption was lower, the effect size and cost-effectiveness of social-norming dropped considerably (Andor *et al.*, 2020). This highlights the importance of considering baseline behaviors and cultural differences when implementing interventions in different contexts. Notably, the effectiveness of social norming can vary even when the behavior being targeted serves the same social purpose. For instance, while both the reduction of household energy consumption and carbon offsetting fall under the broader domain of conservation behaviors, they involve different psychological considerations and barriers. When the social-norming technique, successful in reducing household energy consumption, was applied to carbon offsetting, it did not appear to drive positive behavioral change (Carattini *et al.*, 2024). This can be attributed to the lack of a widespread descriptive norm for carbon offsetting, unlike the established norm to reduce household energy consumption. So, even though these behaviors all work toward the same goal of conservation, the effectiveness of social norming as an intervention may vary due to the distinct nature of each behavior and the psychological factors at play.

Reaching people at the right time when they are receptive can also be tricky. While it may be easy to find the right timing to remind people to wash their hands (e.g., after using the toilet), determining the optimal timing for the ‘Teaching at the Right Level’ (TaRL) pedagogical approach has proven challenging (Banerjee *et al.*, 2017). TaRL worked well in summer camps but faced resistance from teachers and parents when implemented during the regular school year due to their emphasis on covering the grade-level curriculum.

The ‘Timely’ principle goes beyond just the ‘when’ aspect of an intervention; it emphasizes delivering interventions when people are most receptive. Receptiveness can be influenced by factors such as an individual’s stage of life, their decision-making process, or their level of interest in the topic at a given time. Policymakers often cannot individualize the timing for each recipient and instead deliver interventions as a one-time initiative. In such situations, the focus shifts to identifying and targeting the most receptive segment of the population who happen to be at the right timing at the time of intervention delivery. Therefore, at the implementation level, the ‘timely’ principle may also involve segmentation and targeting considerations.

The backfiring of the attractive family-future benefit discussed earlier (Shah *et al.*, 2023) might be the case of an intervention that hit the group below the age of 28 at the wrong time (before they had a family). Carattini *et al.*’s (2024) study on increasing interest in peer-to-peer solar contracts illustrates the intricate relationship between timing and audience targeting. The researchers launched a Facebook campaign aimed at promoting these contracts by showcasing shareable green reports that made pro-environmental behaviors more socially visible. Initially, the campaign effectively engaged the target audience on Facebook, increasing click-through interest by up to 30%. However, as the campaign continued, it began to reach a broader and less interested audience, leading to a decline in effectiveness. This illustrates how the impact of the same intervention can vary over time as it reaches different segments of the population.

While we may be able to conceptually describe an intervention technique (be it simplification, social norming, or message framing) and follow frameworks as guiding principles, the ‘nuances’ in the context can often interact with an intervention and change the final outcome. These examples highlight the pitfalls of simple frameworks (see also Soman, 2017) as well as the importance of considering contextual factors, cultural differences, heterogeneity, and timing when designing and implementing behavioral interventions.

We do not expect media reports to cover and discuss such nuances in detail, as it may be impractical or even impossible to do so, but the omission of even the most basic details on participant characteristics, the population that was targeted by the trial, the physical environment, and the timing of the intervention may lead practitioners to overestimate the generalizability of research findings. As previously discussed, the consequences of low-fidelity media reporting can have far-reaching impacts on the political economy of science. Given the importance of the media reporting fidelity, we introduce a set of rubrics to evaluate how the media presents behavioral research findings. These rubrics aim to assess the accuracy, completeness, and contextual relevance of media reports. Next, we discuss the considerations involved in developing these rubrics.

Rubrics development

We developed rubrics to assess the fidelity of media reports by integrating two established instruments: (a) The Consolidated Standards of Reporting Trial (CONSORT; Schulz *et al.*, 2010) for measuring objective accuracy, and (b) Chang’s (2015) typology of scientific perception of research news for assessing subjective accuracy.

The CONSORT Statement is a set of recommendations for reporting randomized controlled trials (RCTs) that standardizes the way scientists report research findings (Schulz *et al.*, 2010). It emphasizes objective and transparent reporting of trials, facilitating the scientific community’s evaluation and interpretation of research designs. We use the CONSORT Statement for developing items that assess objective accuracy, including details such as how, when, and where an intervention was conducted and to whom the intervention applied.

Chang’s (2015) typology of scientific perception of research news was developed to look into how often health research is misrepresented in the news by doing things like overemphasizing on how unique a study is and overgeneralizing findings. This typology served as the basis for creating items in our rubrics that assess subjective accuracy, including the interpretations of research findings and their implications.

In addition to these established instruments, we incorporated insights from work completed by members of a large international consortium of behavioral scientists and practitioners called the Behaviorally Informed Organizations partnership (BI.Org) (Soman and Yeung, 2021; Mažar and Soman, 2022; Soman, 2024). The resulting rubrics comprised 10 key dimensions for evaluating the fidelity of media reports:

Evidence type: correlational versus causal

Mistaking correlational evidence as causal can lead to costly consequences, including financial losses or loss of life. For example, marketers’ belief that higher advertising

spending drives platform sales has resulted in a waste of approximately USD 50 million per year on eBay (Luca and Bazerman, 2021). In healthcare, the unwarranted conviction linking vaccination and autism has fueled potent anti-vaccination sentiments, leading to strong resistance to immunization, even during urgent times such as the COVID-19 pandemic.

The first dimension of our rubrics assesses the fidelity of media articles in presenting the type of evidence as correlational versus causal. It was not uncommon for media reports to misrepresent the relationship (correlational vs. causal) between variables. For example, a third of UK health research press releases make causal claims based on correlational evidence, give exaggerated advice, or extrapolate animal research to humans (Sumner *et al.*, 2014). About half of the health sections of US, UK, and Canada editions of Google News from July 2013 to January 2014 claimed causal effects despite non-randomized study designs (Haneef *et al.*, 2015). Given the prevalence of mistaking correlational evidence as causal, our rubrics assess the fidelity of media articles in presenting the type of evidence as correlational versus causal.

The attitude/belief–intention–behavior hierarchy

When the original research measures outcomes that are attitudes/beliefs or intentions instead of actual behaviors, the findings have limited practical applicability as they may not reflect real-world behaviors. One of the best examples is perhaps the recurring pattern of unfulfilled New Year's resolutions, where despite sincere intentions, goals often remain unmet. The consequences of the intention–behavior gap extend beyond personal aspirations and influence societal sustainability. For example, while young consumers claim to support ethical consumption, the market share of sustainable organic food and fair-trade products remains disproportionately low, often accounting for less than 1% (Vermeir and Verbeke, 2006).

Given the differences in the practical relevance of attitudinal measures, intentions, and behavioral measures, our rubrics assess the fidelity of media reporting of research outcomes along the attitude/belief–intention–behavior hierarchy.

The use of hypothetical scenarios as the sole basis of conclusions

Research in behavioral social science and cognitive neuroscience has identified differences in brain activities and behavioral responses when comparing hypothetical versus real-life decision-making scenarios (Camerer and Mobbs, 2017). Research findings from studies based on hypothetical scenarios – where participants were asked to imagine themselves in a hypothetical situation, facing a hypothetical intervention, and/or providing a hypothetical response – may not translate directly to real-life situations. Gandhi *et al.* (2024) evaluated 20 pre-registered experiments involving over 15,000 participants and found that hypothetical interventions often yield misleading estimates of real behavior change, cautioning against constructing hypothetical scenarios to assess behavioral interventions. Our rubrics assess whether media reports acknowledge it when the original scientific research uses hypothetical scenarios as the sole basis for conclusion.

Participant characteristics

Most research published in leading journals concerning human psychology and behavior has relied on sampling WEIRD (Western, educated, industrialized, rich, and democratic) populations. Statistically, 96% of studies aiming to construct theory use empirical data from participants who come from countries representing a small fraction of the world's population, approximately 12% (Rad *et al.*, 2018). Race, for example, plays a critical role in a wide range of psychological phenomena and behavioral tendencies (Roberts and Rizzo, 2021). From neural activity to overt behavior, from memory to religious cognition, and from auditory and visual processing to executive functioning, race influences how individuals perceive, interpret, and respond to the world around them (Roberts *et al.*, 2020). Similarly, cultures and demographic variables such as age, gender, and socioeconomic status profoundly influence how people think, develop, and navigate the social world (Brislin, 1993; Pinquart and Sörensen, 2000; Deeks *et al.*, 2009). Several examples discussed earlier (Shah *et al.*, 2023, 2024; Carattini *et al.*, 2024) demonstrated the effectiveness of an intervention varies depending on participant characteristics and are not repeated here. These examples illustrate the importance of understanding the target audience for behavioral interventions and highlight the sensitivity of intervention effectiveness to the characteristics of the recipients. With a growing effort to design behavioral interventions for citizens in less developed countries, who fall outside the typical WEIRD populations, this understanding becomes even more vital. Therefore, our rubrics assess the extent to which media reports accurately present information on participant characteristics.

Elements of context

Human behavior does not occur in a vacuum; it adapts to the circumstances in which it occurs. Context is a multidimensional construct that includes factors such as intervention features, timing, physical environment, and social environment (Yang *et al.*, 2023). Changes in these factors influence behaviors even when the recipient population remains the same. When information about the context is not reported in media articles, practitioners cannot make informed judgments about the applicability of an intervention for their own context. Moreover, the absence of such information may lead to the assumption that interventions are universally applicable, increasing the likelihood of their use in untested situations (List, 2022).

Davis *et al.* (2015) reviewed 82 theories of behavior and behavior change across the social and behavioral sciences and emphasized the importance of broadening perspectives. They advocated for moving beyond a sole focus on individual capabilities and motivations when understanding the theoretical underpinnings of behavioral interventions. They encouraged a more holistic approach that considers contextual variables, such as social and environmental factors, to enhance the effectiveness of behaviorally informed interventions and drive meaningful and sustainable changes. If contextual variables are crucial for forming theories about human behavior that can inform the development of more effective behavioral interventions, media reports should also convey this essential information.

Given these considerations, our rubrics evaluate the media's reporting of contextual factors when such information is available in the original scientific research.

Limitations of the research

Sumpter *et al.* (2023) advocate for the inclusion of a ‘limitations’ section in all scientific papers. It reveals weaknesses in research designs that can impact interpretations and outcomes, guide readers in understanding the potential shortcomings of the original research, and inform methods to mitigate them (Ross and Bibler Zaidi, 2019). Therefore, our rubrics evaluate how faithfully media reports state the limitations as discussed in the original research.

Broad and specific domains of application

Media reports may make inferences and recommendations that extend or generalize beyond the scope of the original research. We have identified three types of (un)faithfulness in the depiction of broad and specific domains in media reports:

1. Cross-domain generalization: This occurs when media reports extrapolate findings from one broad domain to another. For instance, an intervention for promoting physical well-being might be suggested as a solution for improving financial well-being, or an intervention aimed at increasing student engagement in the classroom as a strategy to enhance employee engagement in the workplace.
2. Subdomain generalization: This occurs when media reports discuss the same broad domain as the original study but introduce variables from a different specific (sub)domain. For example, suggesting that a nudge designed to increase participation in weight loss programs, as described in the original research, can also be applied to encourage vaccination uptake. Despite falling under the broad domain of ‘health,’ participation in these health behaviors could be driven by distinct psychological and contextual factors.
3. Overgeneralization: This occurs when media reports broaden the practical relevance of the original research. In other words, the broad domain covered in the original research is only a subset of what the media presents. For example, the original search focuses on a behavioral solution for a specific health issue, but the media report portrays it as a universal ‘cure-all’ solution for various health problems.

It is important to note that our goal is not to judge the severity of different types of generalizations but rather to differentiate among them to uncover nuances. As List (2020) points out, the generalizability of an intervention often depends on whether the new intervention program retains the ‘secret sauce’ of the original one, is implemented in an environment that the theory considers ‘closely exchangeable,’ and whether the participant-specific variables and situational factors theorized as relevant to the original intervention’s success are present in the new setting. Therefore, a ‘subdomain generalization’ may not necessarily pose a lower uncertainty of generalizability than, say, ‘overgeneralization,’ depending on the specific context and factors involved. Nevertheless, by distinguishing between cross-domain generalization, subdomain generalization, and overgeneralization, we can better study how media reports might extend or misrepresent original research findings. This understanding aids researchers,

journalists, and readers alike in evaluating the accuracy and appropriateness of generalizations made in media reports.

Independent and dependent variables

Certain variables may be conceptually related, but they are not interchangeable. Treating one variable as if it were another can lead to misguided intervention strategies. In customer relationship management, the interplay between customer satisfaction and customer loyalty exemplifies this phenomenon. While customer satisfaction has a positive impact on customer loyalty (Fornell *et al.*, 1996), interventions aimed at boosting customer satisfaction (constructed as experience and consumption utility) may not necessarily increase customer loyalty (characterized as decision or choice utility) (Auh and Johnson, 2005). Similarly, in education, learning and performance are often mixed up. Research shows that considerable learning can occur without any immediate performance enhancement, and substantial changes in performance often fail to translate into lasting learning (Soderstrom and Bjork, 2015).

Given the conceptual connection between different concepts, media reporters may intentionally or unintentionally conflate the meanings of one variable with another for a better narrative fit. An accurate and complete understanding of the variables manipulated and measured in the original research is fundamental to developing any behaviorally informed intervention. Therefore, our rubrics assess the representation of independent and dependent variables in media reports.

Ground for recommendations

Media reporters often derive actionable recommendations from scientific research, but these recommendations may or may not be fully grounded in evidence. Practitioners must discern whether recommendations are evidence-based or merely plausible yet untested. This knowledge helps practitioners calibrate their confidence levels when developing behavioral interventions based on media-reported recommendations. To aid in this process, our rubrics assess the sources of recommendations provided by media reports.

Presentation of personal opinions

The distinction between facts and opinions, as outlined in Chapter 56 of *The News Manual* (Ingram, 2024), is crucial in media practice. Facts are statements regarded as true or confirmed, while opinions arise from individual interpretations of factual evidence. The guide emphasizes the need for media reporters to clearly differentiate between these two types of information for their readers. Furthermore, it stresses the importance of attributing opinions and any facts lacking widespread validation. While *The News Manual* promotes good practices, the extent to which reporters follow these principles may vary. Therefore, our rubrics include an examination of how personal opinions are clearly labeled as such when they are presented.

Table 1 displays a list of the fidelity dimensions for which we constructed rubrics. The detailed rubrics can be found in Appendix A.

Table 1. Dimensions of fidelity of media report

Dimensions of fidelity	Explanation
Evidence type: correlational vs. causal	Does the media report accurately and completely mention the relationship between dependent and independent variables as correlational vs. causal, as described in the research article?
The attitude/belief–intention–behavior hierarchy	Does the media report accurately represent the nature of the dependent variable (i.e., attitude/belief, intention, or behavior), as described in the research article?
Using hypothetical scenarios as the sole basis of conclusion	When the research article draws conclusions solely from hypothetical scenarios, does the media report acknowledge it?
Participant characteristics	Does the media report accurately and completely mention each of the participant characteristics described in the research article?
Elements of context	Does the media report accurately and completely mention each of the elements of context described in the research article?
Limitations of the research	How does the media report state the limitations of the research, as discussed in the original research article?
Broad and specific domains of application	Are the broad and specific domains described in the media report consistent with those in the research article?
Independent and dependent variables	Does the media report faithfully represent the independent and dependent variables described in the research article?
Ground for recommendations	Are the recommendations made by the media reporter based on evidence from the research article?
Presentation of personal opinions	How does the media reporter present personal opinions, including their own and those of the research article authors?

Note: Please refer to Appendix A for the exact wordings used in the rubrics.

Fidelity issues in media report with illustrative examples

We used the rubrics to evaluate 68 media reports covering 11 behavioral science studies published in *Nature Human Behaviour* from 2017 to 2021. The full details of the selection process can be found in Appendix B, and the names of these articles and their media coverage are listed in Appendix C. To ensure the quality and reliability of our assessment, the coders who evaluated these research articles and media reports were social science Ph.D. students in at least their 2nd year at the University of Toronto.¹

The analysis provided preliminary results of media reporting patterns across the ten dimensions listed in Table 1, some of which are particularly prone to fidelity issues.

¹These coders had a fundamental understanding of various types of variables, including independent and dependent variables, as well as contextual factors and participant characteristics. We specifically sought coders with a good grasp of different research methodologies, such as RCTs and observational studies. At the time of coding, all but one coder had already passed their Comprehensive Exams, demonstrating their expertise in the field. Given the complexity of the evaluation process and the need for nuanced understanding of the research and media reports, we are unaware of any non-human or AI-based systems that could do the coding. As a result, the coders had to manually evaluate all the elements in the rubric, resulting in a resource-intensive process. Perhaps suitable AI systems might make the coding process easier and therefore more scalable.

Table 2. Media reports on relationships between variables in our sample of NHB articles

	Reported by media as	
	Causal	Correlational
Scientific evidence: causal (RCT-based)* ⁺	11	2
Scientific evidence: causal (non-RCT-based)* ⁺	18	0
Scientific evidence: correlational	6	15

*Our rubrics distinguished between causation derived from traditional RCTs and causation derived from advanced statistical techniques (non-RCT based). We considered both forms of causal evidence when compiling the statistics for this commentary.

⁺Research is considered to have established causal evidence if its empirical package includes at least one study demonstrating causation.

Table 3. Media reports on dependent variables along the attitude/belief–intention–behavior hierarchy in our sample of NHB articles

	Reported by media as		
	Attitude/belief	Intention	Behavior
Scientific evidence: attitude/belief	2	0	0
Scientific evidence: intention	0	9	0
Scientific evidence: behavior ⁺⁺	0	0	42

⁺⁺Research is considered to have established behavioral evidence if its empirical package includes at least one study demonstrating the behavioral outcome of interest.

Evidence type: correlational vs. causal

Each media report was assessed by a pair of coders, and only those assessments that achieved inter-coder reliability (i.e., consistent assessments between the coders in each pair) were accepted. Inter-coder reliability was achieved for 52 out of the 68 media reports. Among these 52 media reports, 44 accurately described the nature of evidence, yielding an overall fidelity rate of 85% in our sample. However, upon closer examination of the 21 media reports covering research articles with correlational scientific evidence, only 15 correctly identified it as such, while six erroneously described it as causal. As a result, the fidelity rate for accurately identifying correlational evidence in our sample was 71.4%. A detailed breakdown is shown in [Table 2](#).

The attitude/belief–intention–behavior hierarchy

Inter-coder reliability was achieved for 53 out of the 68 media reports. All 53 media reports accurately reported scientific evidence at all levels of the attitude/belief–intention–behavior hierarchy ([Table 3](#); the diagonal shows the number of times the reporting was correct).

The use of hypothetical scenarios as the sole basis of conclusion

In this part of the analysis, we focused on scientific research that based its conclusions solely on hypothetical scenarios and assessed the extent to which media reports accurately reflected this basis. We structured our evaluation into two phases.

First, we identified research projects that exclusively used hypothetical scenarios in their empirical work. Nine out of the 11 research projects in our sample achieved inter-coder reliability. Out of these, three projects relied solely on hypothetical scenarios while the remaining six incorporated field evidence.

Subsequently, we analyzed the media coverage of the three research projects whose conclusions were based entirely on hypothetical scenarios. A total of 22 media reports discussed these projects. Inter-coder reliability was achieved for 18 of these media reports regarding whether the reports acknowledged the exclusive use of hypothetical scenarios in the original research. Of these 18 media reports, only 5 (28%) explicitly mentioned the original research's reliance on hypothetical scenarios. The other 13 reports (72%) did not acknowledge so.

Participant characteristics and elements of context

We evaluated how nine common participant characteristics were described in the original research and presented in their corresponding media reports: age, gender, ethnicity, religion, occupation, education, income, health condition, and marital status. Additionally, we conducted the same assessments for five contextual elements: timing, physical environment, interface, social & cultural environment, and choice architecture. For all these dimensions, high levels of inter-coder reliability were achieved; however, since the exact assessment varied across each dimension, we report such numbers in the Appendices (Appendix D1 and Appendix D2).

Table 4 shows the fidelity of the media reports in our sample in presenting participant characteristics and elements of context when such information was presented in the original scientific research articles. The fidelity rate is the percentage of media reports that faithfully presented the information when it was available (i.e., presented) in the corresponding scientific research article. The media's presentation of participant attributes appears to be somewhat sparse and varied across different attributes. For example, only 40% of media reports in our sample faithfully mentioned participant occupation when occupation was described in the original research. Even more notably, only 14% reported the age of the participants. Other details regarding gender, religion, ethnicity, education, income, and health condition of the research participants were often omitted or inaccurately reported.

In terms of contextual information, only 34% of media reports in our sample faithfully mentioned the physical settings where the original research took place. Regarding the reporting of timing, interface, and social and cultural environment of the original research, the rate of accurate reporting dropped to 29%, 23%, and 11%, respectively. Faithful reporting of choice architecture was even lower, at just 3%.

Limitations of the research

All 11 NHB research articles in our sample discussed the limitations of their research, suggesting that all the 68 media reports had access to this information. Inter-coder reliability was achieved for 53 media reports, of which only two (4%) accurately and completely mentioned the limitations described in the corresponding research articles. The remaining 51 (96%) reports did not appear to include any mention of the research limitations.

Table 4. Media reports on participant characteristics and elements of context in our sample of NHB articles

Participant characteristics	Number of media reports that have access to information on the target characteristic*	Number of media reports that faithfully present information on the target characteristic	Fidelity rate (percent)
Age	29	4	14%
Gender	29	0	0%
Ethnicity	23	0	0%
Religion	22	0	0%
Occupation	25	10	40%
Education	28	0	0%
Income	26	0	0%
Health condition	7	0	0%
Marital status	0	0	NA
Elements of context	Number of media reports that have access to information on the target element**	Number of media reports that faithfully present information on the target element	Fidelity rate (percent)
Timing <i>The right moment (rather than a specific point in the timeline)</i>	48	14	29%
Physical environment <i>Characteristics of the physical environment where experimentation takes place</i>	38	13	34%
Interface <i>The format in which experimentation takes place</i>	57	13	23%
Social and cultural environment <i>Influences from the 'people' side</i>	46	5	11%
Choice architecture <i>How choices are presented</i>	29	1	3%

*This is the case when the information on participant characteristics was explicitly provided in the corresponding original research article.

**This is the case when the information on elements of context was explicitly provided in the corresponding original research article.

Independent and dependent variables

Of the 48 media reports with inter-coder reliability in their assessment of independent variables, 43 (90%) described the same ones as the original research, while 5

(10%) described different independent variables. In terms of the media reporting of dependent variables, among the 48 media reports with intercoder reliability, 40 (83%) portrayed the same ones as the original research, while 8 (17%) deviated.

Ground for recommendations

Out of the 68 media reports, 38 achieved inter-coder reliability. Of these 38 media reports, only 23 (61%) appeared to give recommendations directly backed by evidence presented in the original research article. The remaining 15 (39%) inferred recommendations from research findings; in other words, these recommendations were not explicitly tested in the original research.

Presentation of personal opinions

The coders evaluated the presentation of personal opinions in media reports, categorizing them into four types. Some categories were mutually exclusive (e.g., the first and the second categories in Table 5 below), while others could co-occur in the same report. Additionally, a single report could exhibit multiple instances across these categories (e.g., a report might simultaneously offer unacknowledged personal opinions of the reporter and misrepresent the opinions of research authors as facts).

Assessing how media writers present their personal opinions involves considerable subjective judgment. We achieved inter-coder reliability for 32 out of the 68 media articles analyzed. A total of 33 instances were identified across these categories: (a) In 19 media reports, reporters offered their own opinions and labeled them as such; (b) in eight media reports, reporters offered their own opinions without acknowledging so; (c) in five media reports, reporters presented opinions discussed by research article authors as scientific facts; and (d) in one media report, reporters mixed their own opinions with those of research article authors'. Overall, slightly more than half of the media reporters in our sample clearly stated their opinions as their own.

Broad and specific domains of application

Our rubrics required coders to determine and label the domain(s) of application covered in research articles and media reports, respectively. The primary purpose was to guide the coders' subsequent judgment on whether media reporters altered the domain(s) of application in a way that deviated from the original research. Therefore, our analysis focused on evaluating the consistency between the domains covered in

Table 5. Media reports on presentation of personal opinions in our sample of NHB articles

Media reporters	Number of occurrences
1. Offered their own opinions and labeled as such	19
2. Offered their own opinions without acknowledgment	8*
3. Presented research article authors' opinions as facts	5*
4. Mixed their own opinions with research article authors'	1

*In one media report, the reporters offered their personal opinions without acknowledgment and presented the opinions of research article authors as facts.

research articles and those discussed in their media reports, rather than on how these domains were labeled.²

Two noteworthy observations emerge regarding media reporting on broad and specific domains of application in our sample of behavioral science research: (a) Media reports generally stay faithful to the representation of broad domains covered in research articles. Inter-coder reliability was achieved in assessing broad domains for 51 out of the 68 media reports, with all 51 (100%) matching the broad domains described in the research article. (b) For specific domains, inter-coder reliability was achieved in assessing specific domains for 52 out of the 68 media reports. Of these 52 reports, 41 (79%) matched the specific domains described in the research article, while the remaining 11 (21%) deviated,³ exhibiting cross-domain generalization, subdomain generalization, or overgeneralization.

Summary of observations

In conveying the nature of evidence, misreporting correlational evidence as causal appears to be prevalent. Additionally, when evidence is based solely on hypothetical scenarios, its hypothetical nature is often not acknowledged, according to our initial evidence.

Media reports in our sample appear to omit details about research participants and contexts. While age and occupation of participants are occasionally mentioned, other key attributes such as gender, ethnicity, religion, education, income, and health condition are rarely reported, even when such attributes were presented in the original research. Similarly, the media's reporting on contextual dimensions also falls short. The media's tendency to omit the reporting of participant characteristics and context hinders practitioners' ability to realize how these factors may influence research outcomes. Moreover, practitioners might mistakenly view behavioral interventions as universally effective (List, 2022) and neglect the importance of testing interventions for specific contexts.

Fidelity rates in reporting research limitations and personal opinions also appeared to be low. Only 4% of media reports in our sample accurately and completely mention the limitations described in the original research articles. This inadequate coverage can lead practitioners to underestimate the limitations of the interventions being tested and prevent them from properly calibrating their confidence levels when applying and scaling up these interventions. Additionally, the failure to distinguish between the opinions of media reporters, the opinions of research article authors, and scientific findings impedes practitioners' ability to discern subjective views from objective facts without referring to the original research articles.

²The labels were not further analyzed due to significant variations in the coders' chosen wordings. For instance, research on using technology to improve health could be labeled as 'health,' 'technology,' 'e-health,' or 'health technology.' Standardizing the choice of labels across coders without an established taxonomy would be challenging.

³Inter-coder reliability was achieved among these 11 media reports at the level of indicating low-fidelity reporting. However, the exact form of low-fidelity, be it cross-domain generalization, subdomain generalization, or overgeneralization may be judged differently among coders.

Table 6. Decision-support tool: a checklist for evaluating media reports of behavioral science research

Evidence type and quality	
Does the media report indicate whether the evidence is causal or correlational?	Y/N
Does the media report indicate if the original research drew conclusions based entirely on hypothetical scenarios?	Y/N
Population characteristics and contextual elements	
Does the media report mention key characteristics of the study population, such as demographics and psychographics?	Y/N
Does the media report describe elements of the study context, such as timing, physical environment, and social and cultural environment?	Y/N
Research limitations and constraints	
Does the media report acknowledge any limitations of the original research?	Y/N
Opinion attribution	
Does the media report explicitly attribute opinions to specific individuals or sources?	Y/N

Fidelity issues in media reporting and ways to handle them

Our preliminary analysis uncovers two broad types of fidelity issues. The first type is the omission of information (e.g., on participant characteristics, contextual elements, and limitations of research); the second type is the inaccuracy of information (e.g., presenting correlational evidence as causal, presenting conclusions drawn from hypothetical scenarios as field findings, making unwarranted generalizations, presenting personal opinions as facts). To address these issues, we propose a checklist that highlights frequently overlooked dimensions and common sources of inaccuracy in media reporting of behavioral science research.

The checklist is designed to enhance practitioners' awareness, ensuring they grasp critical aspects of the original research they might otherwise overlook and highlighting areas where inaccurate information tends to arise. This tool is particularly suitable for addressing the first type of issue – omission of information. For the second type of issue, inaccuracy of information, while the checklist can raise awareness, identifying inaccuracies will require further effort to compare media reports with the original scientific research.

A practitioner's checklist for assessing media coverage of behavioral science research

We present a checklist in [Table 6](#) as a decision-support tool for practitioners. This tool directs practitioners' attention to often overlooked dimensions and common sources of inaccuracy in media-reported behavioral science research. It provides a systematic approach to help practitioners critically evaluate such findings as reported through media channels.

Answering NO to any of these questions does not immediately disqualify a media report as a high-fidelity source. However, it does suggest that the media report may lack important details from the original research, prompting practitioners to be cautious and possibly consult the original research for a complete understanding.

Tailoring checklist to research stage

The external validity of our checklist may vary depending on the stage of the original scientific research. We suggest adaptations of the checklist based on List's (2020) research hierarchy framework. These changes would be made to fit the needs of different research stages with varying levels of readiness for scaling.

Wave 1 research

Wave 1 research often involves the use of controlled laboratory experiments or proof-of-concept studies. These experiments are conducted in a highly controlled environment where the researcher can manipulate variables precisely and observe outcomes in a clean and controlled setting. Although Wave 1 research can offer compelling proof-of-concept evidence, its external validity is often limited. Practitioners are advised to treat media reports of Wave 1 research as preliminary and be highly cautious about generalizing these findings to real-world settings, even if media reports meet many criteria in the checklist in [Table 6](#).

Wave 2 research

Wave 2 research typically evaluates interventions in more heterogeneous populations and settings. It examines boundary conditions and aims to replicate results in real-world settings. When reading media reports of Wave 2 research, practitioners are encouraged to look for detailed information on sample characteristics, contextual factors, and the robustness of results across different conditions. The absence of this information should prompt a consultation of the original research article to assess the applicability of the findings to specific contexts.

Wave 3 research

Wave 3 research directly informs practitioners' scaling efforts. Researchers at this stage clearly delineate 'negotiables' (design choices that can vary without affecting key outcomes) and 'non-negotiables' (features essential for maintaining impact) to guide implementation at scale. When reading media reports of Wave 3 findings, practitioners might expect comprehensive discussion of scale-up considerations, including any evidence on the generalizability of results to policy-relevant populations and contexts. If a media report does not provide such information, practitioners are encouraged to refer to the original research article for guidance in their implementation decisions.

Notably, these adaptations emphasize the importance of researchers' clearly labeling their research according to the stage it maps to within the research hierarchy, as List (2020) suggested. Media reporters are also encouraged to do the same. This approach ensures that practitioners can make well-informed decisions based on the readiness of research findings for practical application, as delineated in List's (2020) research hierarchy.

Using the checklist: three key usage scenarios

The checklist can be used in various scenarios in an organization to guide the use of behavioral science research.

Assessment tool

Practitioners can use the checklist to assess media reports, identifying gaps and inconsistencies in the presented information. This checklist helps in making informed decisions about the applicability and reliability of the findings.

Customization for context

Adapting the checklist to the specific organizational context, target population, and intervention goals optimizes its usefulness. Integrating the checklist as an essential tool within the broader knowledge management framework embeds these practices into the organization's culture, creating a repeatable formula for achieving desired change outcomes.

Knowledge repository

Documenting evaluation results and building a repository of institutional knowledge that informs future intervention work is crucial. Regularly reviewing and updating the customized checklist ensures its continued relevance and effectiveness as a knowledge management tool.

Conclusion

Media reports play a crucial role in disseminating behavioral science research findings to an ever-growing body of behavioral practitioners, yet our analysis reveals potential fidelity issues that can hinder the effective translation of research into practice. Important contextual information and participant characteristics may be left out, correlational evidence may be mistaken for causal evidence, and the hypothetical nature of some research is not acknowledged. This can lead to an incomplete and sometimes misleading picture of the original research. Furthermore, the inability to distinguish between evidence-based findings and subjective opinions can potentially misguide the scaling of (untested) interventions in the field.

Notably, our findings are based on a limited sample, so further assessments are required to confirm the generalizability of these insights. Alternatively, we have shared the rubrics and the considerations we had in developing them, so that readers, if they wish to conduct a similar analysis for their specific field of study, could use ours or adjust them to cater to their specific needs.

We propose a checklist that serves as a tool for evaluating media reports. The checklist helps practitioners identify information gaps and adjust their level of confidence in the findings' usefulness and accuracy in their own situations.

The onus of ensuring faithful reporting does not rest solely on the practitioners. Researchers must strive for clarity and transparency in their communication, explicitly labeling their research stage and its readiness for practical application (List, 2020). Media reporters, in turn, bear the responsibility of accurately conveying the nuances and limitations of the studies they cover, resisting the temptation to oversimplify for the sake of a compelling narrative. Only through a concerted effort by all stakeholders – researchers, media reporters, and practitioners – can we create a robust and reliable pipeline for translating behavioral science insights into real-world impact. By fostering

a culture of critical evaluation, contextual awareness, and evidence-based decision-making, we can unlock the potential of behavioral science to drive positive change in organizations and society at large.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/bpp.2025.5>.

Funding statement. This research was funded by a Social Sciences and Humanities Research Council of Canada (SSHRC) Partnership Grant 895-2019-1011 (Behaviorally Informed Organizations) to Soman.

Competing interests. The authors declare no competing interests.

Author contributions. Jingqi Yu, Catherine Yeung, and Dilip Soman conceived the project and developed the rubrics. Jingqi Yu oversaw the coding and analyzed coding results. All the authors contributed to early drafts of the manuscript. Jingqi Yu and Catherine Yeung wrote the final manuscript. Dilip Soman supervised the project.

Informed consent. This article does not contain any studies with human participants performed by any of the authors.

References

- Allcott, H. (2011), 'Social norms and energy conservation', *Journal of Public Economics*, **95**(9–10): 1082–1095.
- Andor, M. A., A. Gerster, J. Peters and C. M. Schmidt (2020), 'Social norms and energy conservation beyond the US', *Journal of Environmental Economics and Management*, **103**: 102351.
- Auh, S. and M. D. Johnson (2005), 'Compatibility effects in evaluations of satisfaction and loyalty', *Journal of Economic Psychology*, **26**(1): 35–57.
- Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukerji, M. Walton (2017), 'From proof of concept to scalable policies: challenges and solutions, with an application', *Journal of Economic Perspectives*, **31**(4): 73–102.
- Behavioural Insights Team (2014), 'EAST: four simple ways to apply behavioural insights'. The Behavioural Insights Team. Accessed on December 1, 2024. <https://www.bi.team/publications/east-four-simple-ways-to-apply-behavioural-insights>
- Briscese, G. and J. A. List (2024), 'Toward an understanding of the political economy of using field experiments in policymaking', NBER Working Paper Series, No. 33239, December 2024, <http://www.nber.org/papers/w33239>
- Brislin, R. (1993), *Understanding Culture's Influence on Behavior*, Forth Worth: Harcourt Brace Jovanovich College Publishers.
- Camerer, C. and D. Mobbs (2017), 'Differences in behavior and brain activity during hypothetical and real choices', *Trends in Cognitive Sciences*, **21**(1): 46–56.
- Carattini, S., K. Gillingham, X. Meng and E. Yoeli (2024), 'Peer-to-peer solar and social rewards: evidence from a field experiment', *Journal of Economic Behavior & Organization*, **219**: 340–370.
- Chang, C. (2015), 'Inaccuracy in health research news: a typology and predictions of scientists' perceptions of the accuracy of research news', *Journal of Health Communication*, **20**(2): 177–186.
- Crowley, D. M., J. T. Scott, E. C. Long, L. Green, A. Israel, L. Supplee, C. Giray (2021), 'Lawmakers' use of scientific evidence can be improved', *Proceedings of the National Academy of Sciences*, **118**(9): e2012955118.
- Davis, R., R. Campbell, Z. Hildon, L. Hobbs and S. Michie (2015), 'Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review', *Health Psychology Review*, **9**(3): 323–344.
- Deeks, A., C. Lombard, J. Michelmores and H. Teede (2009), 'The effects of gender and age on health related behaviors', *BMC Public Health*, **9**: 1–8.
- Fornell, C., M. D. Johnson, E. W. Anderson, J. Cha and B. E. Bryant (1996), 'American customer satisfaction index: nature, purpose, and findings', *Journal of Marketing*, **60**(4): 7–18.

- Gandhi, L., A. Kiyawat, C. Camerer and D. J. Watts (2024), 'Hypothetical nudges provide misleading estimates of real behavior change', *Academy of Management Proceedings*, **2024**
- Haneef, R., C. Lazarus, P. Ravaud, A. Yavchitz and I. Boutron (2015), 'Interpretation of results of studies evaluating an intervention highlighted in Google health news: a cross-sectional study of news', *PLoS One*, **10**(10): e0140889.
- Hubble, C., and C. Varazzani (2023). *Mapping the global behavioural insights community* Observatory of Public Sector Innovation. <https://oecd-opsi.org/blog/mapping-behavioural-insights/> (Accessed 1 December 2024)
- Hjort, J., M. Diana, R. Gautam and F. S. Juan (2021), 'How research affects policy: experimental evidence from 2,150 Brazilian municipalities', *American Economic Review*, **111**(5): 1442–1480.
- Ingram, D. (2024). The news manual: a professional source for journalists and the media Retrieved from: https://www.thenewsmanual.net/Manuals%20Volume%203/volume3_56.htm (Accessed 1 June 2024)
- Kuehn, S. A., and A. Lingwall (2016), *The Basics of Media Writing: A Strategic Approach*, SAGE Publications, CQ Press
- List, J. A. (2020), Non Est Disputandum De Generalizability? A glimpse into the external validity trial. NBER Working Paper No. w27535 Available at SSRN: <https://ssrn.com/abstract=3658829> (Accessed September 2021).
- List, J. A. (2022), *The Voltage Effect: How to Make Good Ideas Great and Great Ideas Scale* Crown Currency.
- Luca, M. and M. H. Bazerman (2021), *The Power of Experiments: Decision Making in a Data-driven World*, MIT Press
- Mažar, N., and D. Soman (Eds.) (2022), *Behavioral Science in the Wild*, Toronto: University of Toronto Press
- Pinquart, M. and S. Sörensen (2000), 'Influences of socioeconomic status, social network, and competence on subjective well-being in later life: a meta-analysis', *Psychology and Aging*, **15**(2): 187.
- Rad, M. S., A. J. Martingano and J. Ginges (2018), 'Toward a psychology of *Homo sapiens*: making psychological science more representative of the human population', *Proceedings of the National Academy of Sciences*, **115**(45): 11401–11405.
- Roberts, S. O., C. Bareket-Shavit, F. A. Dollins, P. D. Goldie and E. Mortenson (2020), 'Racial inequality in psychological research: trends of the past and recommendations for the future', *Perspectives on Psychological Science*, **15**(6): 1295–1309.
- Roberts, S. O. and M. T. Rizzo (2021), 'The psychology of American racism', *American Psychologist*, **76**(3): 475.
- Ross, P. T. and N. L. Bibler Zaidi (2019), 'Limited by our limitations', *Perspectives on Medical Education*, **8**: 261–264.
- Schultz, P. W., J. M. Nolan, R. B. Cialdini, N. J. Goldstein and V. Griskevicius (2007), 'The constructive, destructive, and reconstructive power of social norms', *Psychological Science*, **18**(5): 429–434.
- Schulz, K. F., D. G. Altman and D. Moher (2010), 'CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials', *Journal of Pharmacology & Pharmacotherapeutics*, **1**(2): 100–107.
- Shah, A. M., M. Osborne, J. Lefkowitz Kalter, A. Fertig, A. Fishbane and D. Soman (2023), 'Identifying heterogeneity using recursive partitioning: evidence from SMS nudges encouraging voluntary retirement savings in Mexico', *PNAS Nexus*, **2**(5): pgad058.
- Shah, A. M., M. Osborne, D. Soman, J. Lefkowitz Kalter, A. Fertig, and A. Fishbane (2024), 'The beneficent and maleficent effects of simplification: evidence of fluency amplification effects on retirement savings contributions', *Journal of Consumer Research*; ucae067
- Soderstrom, N. C. and R. A. Bjork (2015), 'Learning versus performance: an integrative review', *Perspectives on Psychological Science*, **10**(2): 176–199.
- Soman, D. (2017). The elegant simplicity (and potential pitfalls) of simple frameworks. In OPRE Report 2017-23, Nudging Change in Human Services, Office of Planning, Research and Evaluation, U.S. Administration for Children and Families.
- Soman, D. (Ed.) (2024), *What Works, What Doesn't (And When): Case Studies in Applied Behavioral Science*, Toronto: University of Toronto Press
- Soman, D., and C. Yeung (Eds.) (2021), *The Behaviorally Informed Organization*, Toronto: University of Toronto Press
- Stefano, C., R. Dur and J. List (2024), 'Policy evaluation and the causal analysis of public support', *Science*, **386**: 490–492. doi:[10.1126/science.adp7497](https://doi.org/10.1126/science.adp7497)

- Sumner, P., S. Vivian-Griffiths, J. Boivin, A. Williams, C. A. Venetis, A. Davies et al. (2014) ‘The association between exaggeration in health related science news and academic press releases: retrospective observational study’, *British Medical Journal*, **349**: g7015–g7015.
- Sumpter, J. P., T. J. Runnalls, A. C. Johnson and D. Barcelo (2023), ‘A ‘Limitations’ section should be mandatory in all scientific papers’, *Science of the Total Environment*, **857**: 159395.
- Thaler, R. H., and C. R. Sunstein (2021), *Nudge: The Final Edition*, New York: Yale University Press
- Toma, M. and E. Bell (2024), ‘Understanding and increasing policymakers’ sensitivity to program impact’, *Journal of Public Economics*, **234**: 105096. doi:[10.1016/j.jpubeco.2024.105096](https://doi.org/10.1016/j.jpubeco.2024.105096)
- Vermeir, I. and W. Verbeke (2006), ‘Sustainable food consumption: exploring the consumer “attitude–behavioral intention” gap’, *Journal of Agricultural and Environmental Ethics*, **19**: 169–194.
- Yang, S. et al. (2023). The elements of context. Toronto, Canada: Behavioral Economics in Action at Rotman (BEAR) Report series.
- Yu, J. and B. Feng (2024). Where do behavioural practitioners learn about the science. Toronto, Canada: Behavioral Economics in Action at Rotman (BEAR) Report series.