What is wrong with advertising research and how can we fix it?

What is wrong with advertising research and how can we fix it?

Patrick De Pelsmacker

University of Antwerp (Belgium), Faculty of Business and Economics, department of marketing

Abstract

Advertising research is an applied discipline. It should use insights and theories from other disciplines to develop actionable recommendations for contemporary advertising practice, by explaining in a valid way how and why advertising works. I discuss some observations that indicate that part of the current academic advertising literature falls short of this expectation, using information in 330 articles that were published in the International Journal of Advertising (200) and the Journal of Advertising (130) in the period 2016-2019: Lack of realism, no or very little practical relevance, analytical stunt work to study relatively trivial topics, sloppy sampling methods, a lack of novel methods, hardly any research on real behavior, very little qualitative research, relatively little 'integration', and the obsessions and prejudices of the academic bubble. I also provide suggestions to improve academic advertising research, such as the changing the gatekeeper role and habits of editors and reviewers, adapting guidelines for authors, fundamentally altering the review process, for instance by means of preregistration of studies, and changing the criteria to hire and evaluate advertising researchers and professors.

Keywords: advertising research, managerial implications, research methods

Introduction

When was the last time you reviewed a submitted or read a published advertising paper and were curious about the results? These are not my words, but the words of a senior non-cynical and highly regarded advertising scholar. Although his plaint is obviously exaggerated, in my experience it reflects the feeling of many others in the field. The starting point of my reflections in this piece is simple. Advertising research is an applied discipline. It should use insights and theoretical frameworks from various other disciplines to advance advertising practice, by explaining in a valid way how and why it works. There are several observations that seem to indicate that part of the current academic advertising literature falls short of this expectation. In what follows, I will discuss some problematic issues that are indicative of this. However, although gut feeling is a good starting point for forming an opinion, it is not enough. Therefore, I screened 330 articles that were published in the International Journal of Advertising (200) and the Journal of Advertising (130) in the period 2016-2019. I left out editorials, short opinion pieces, and the entire JA 2017 issue 1 on methodology in advertising research. As to the latter, I did not exclude this special issue because of its lack of quality or relevance. On the contrary, it is an excellent and very relevant issue for advertising researchers. However, this issue is of minor relevance for my analysis (the 'counts') since I was interested in the general research practice in adverting research. Throughout the manuscript, I refer to some of the articles in this special issue because they provide solutions and guidelines with respect to some of the research practices I discuss.

Do not expect a rigorous content analysis with Krippendorff's Alpha intercoder reliability measures. The numbers I report are only indicative of the points I want to make. Also, do not read this piece as a frontal attack on authors, reviewers and editors. They mean well, and most of them

are skilled researchers. I am also only going to give general examples and I will refrain from giving examples about specific articles. It is also not my intention to come to a conclusion about which journal does a better job. The numbers I report are across all 330 articles. Here is my list of observations that I checked in these articles: no or very little practical relevance, lack of realism, analytical stunt work to study relatively trivial topics, the inappropriate use of student samples and sloppy sampling methods, a lack of novel methods, hardly any research on real behavioral data, very little qualitative research, relatively little 'integration', and the obsessions and prejudices of the academic bubble. One last note. I am by no means the first or the only one that has thought about these issues. Also consult the excellent contributions by Bergkvist (2020), Taylor (2020) and Royne (2016).

The elephant in the room

Recently, one of my Ph.D. students defended his thesis. The last couple of years he worked part-time in a large advertising agency, while finalizing his thesis. His research work is published in good journals. At his defense, one of his colleagues at the agency asked him in what way his Ph.D. work had been useful for his job at the agency. His answer – with a big smile - was: none of it whatsoever. This is illustrative of the fact that, often, academic advertising research and professional advertising are two distinct worlds, and that both worlds do not seem to perceive this as problematic (read Schultz (2016) for a practitioner vs. academic view of advertising (research)). In my opinion, it is.

One of the problems is the lack of realism in many studies. Stimuli in advertising research have to be professionally made, of high quality, and realistic. Stimulus and context realism become increasingly important due to the constant evolution of advertising and advertising contexts. New advertising formats and contexts such as, for instance, in-game advertising, native advertising, content marketing, and virtual/augmented reality ads, require realistic stimuli that reflect the real world of advertising. For instance, testing the effects of in-game advertising requires that the ad is integrated into a game. The researcher could develop static stimuli ("stills" of the game) in which in-game ads are integrated in different ways, but they would be far away from a real gaming experience, and it is then difficult to validly test the effect of, for instance, interactivity or intrusiveness. Another example is unrealistically long (forced) exposure conditions or very few stimuli that are not embedded in a media context in experiments. This is a big problem for the external validity of advertising research. Think of, for instance, measuring brand recall after a single ad exposure of six seconds for one ad (see, for a discussion, Erfgen et al., 2015). No wonder that advertising practitioners are not keen on learning from advertising researchers.

There is nothing so practical as a good theory (Kurt Lewin). As academics, one of our core tasks is to substantiate our recommendations for advertising practice by means of relevant theory and adequate research methods. However, a major problem is that, often, this academic focus on theory and methods goes at the expense of practical relevance. Looking at the 330 articles, these are the problems. In some cases there is no 'managerial implications' section or paragraph at all. The reader is expected to infer practical relevance from the discussion section. In other cases, the contribution to practice is 'old news', trivial, or very generally stated. And in many cases, the recommendations are not actionable, in the sense that, practically speaking, they are not doable or operational for advertisers.

Into the category of ‘old news’ fall contributions such as ‘we already knew this from previous studies, but now we established that it is also true for Belgium’. Don’t get me wrong. Replication studies are essential for theory building and for solid conclusions about how things work. However, extended replications with an additional angle that clarifies the boundary conditions of a phenomenon are better. In some cases, the contribution to practice is trivial: ‘which color is a better background for a native ad?’ I would say: do not tire yourself finding theories and doing a series of studies to find out something that can easily be solved by means of a simple test. In many cases, the managerial implications are very generally stated: ‘advertisers should take the results of our study into account to segment their audiences in a more relevant way and deliver more persuasive messages to them’. And when practical recommendations are given, they are often not actionable in that an advertiser cannot really put them into practice: ‘advertisers could develop different messages for chronic promotion- and prevention-oriented customers’. Apparently, often, writing a managerial contribution section is perceived as an annoying last-minute job, preceded by the existential insight: this is never going to be of any use to anybody, but let’s try to make the best of it.

Exactly half of the 330 articles I screened do not have a substantial, specifically stated and actionable set of recommendations for advertising practice. Of course, this is my personal judgement. Screening the same articles, others could come up with 40 or 60 percent. But, in fact, it should be zero percent. One important reason for non-actionable recommendations is that one in four articles include individual traits, such as chronic regulatory focus, self-construal, or individualism-collectivism at the individual level. You know the examples: ‘The effect of X on Y: the moderating role of self-construal’. Theoretically relevant as these studies may be, needless to say that it is not easy to use them in advertising practice. Some authors try to solve this problem

by pointing at socio-demographic correlates ('young people and women on average score higher on…'), but that is usually far from sufficient to be practically relevant.

One other reason for this phenomenon may be the way universities hire and evaluate professors and researchers. Most universities only focus on the number of papers and the journals these papers have been published in. Practical relevance of research is most often less or even not considered. I will come back to this in later sections.

Still, many authors do a good job, though, and explain in relevant detail how to make sure their findings can be made actionable in practice.

Give them a big round of applause, ladies and gentlemen!

Ten percent of the articles fall into the category of what I would call stunning methodological stunt work. They fall into two subcategories. The first one is to report, say, 6 studies, each of which developed with a different theoretical framework, eventually leading to relatively obvious or practically irrelevant findings. The second subcategory consists of extremely complicated conceptual and analytical boxes-and-arrows frameworks, employing Structural Equation Modeling or PROCESS models in the high numbers. They come with titles such as: 'the effect of X on Y, serially mediated by M1, M2 and M3: the moderating role of W'. Their 'results' sections are most of the time an illustration of the saying: 'a cat couldn't find her kittens in there', or of the beautiful French expression: 'l'art pour l'art' (art for art's sake). Again, I admire these authors. They are walking libraries and extremely skillful methodologists, but they overdo it. Parsimony and giving precedence to simplicity in conceptual frameworks and analytical methods is an old

and valuable principle, especially in a practical field such as advertising. Look up the 14th century Ockham's razor: 'plurality should not be posited without necessity' (https://www.britannica.com/topic/Occams-razor).

Students are a researcher's best friends

How often do we not come across the sentence 'participants were undergraduate students from a mid-sized South-Eastern university'? Indeed, a substantial part of advertising research is based on student samples from Western countries. Thirty percent of the 330 articles I screened used student samples, the vast majority of them in multi-study articles, in combination with samples of 'real people'. Of course, I did not count studies in which student samples are obviously necessary, such as research on children and teenagers. Advertising research is sometimes labeled 'the science of the undergraduate marketing student' (Jones and Sonner 2001). Some authors feel a bit embarrassed about this, and refer to their sample as 'young millennials, average age 20.4 years'. Most of them do not give any justification for their use of student samples. The ones who do always give similar reasons, such as: 'students are avid users of social media (numbers, numbers…), and are therefore a relevant population in the context of this study'. Another argument that is often used is that a homogeneous population such as students is suitable for testing manipulated effects because, in that way, confounding effects of age, level of education or lifestyle can be avoided (Fleck and Maille 2010).

The relevance of student samples in research is an old debate. There are studies showing that results received in student samples are congruent with the results gained in other populations or

representative samples (e.g., Völckner and Sattler's 2006; Fuchs and Sarstedt 2010). However, a lot of research comparing findings in student samples and samples of 'real people' has found differences in the results obtained from the two samples, and most conclude that students are not good surrogates for adult consumers (for a discussion and overview, see, for instance, Jones and Sonner 2001). Students are, in relevant ways, different from 'ordinary' consumers. They usually have less money, their life style is atypical for the consumption situation of most consumers, and they usually have no or less experience in buying most consumer products. This may easily bias their perceptions, motivations and preferences. (Geuens and De Pelsmacker 2017). Henrich, Heine and Norenzayan (2010, p.29) state that 'people from Western, educated, industrialized, rich and democratic (WEIRD) societies — and particularly American undergraduates — are some of the most psychologically unusual people on Earth'.

However, there are a few cases in which student samples may appropriate. For instance, in studies focusing psychological processes, student samples are adequate if there is evidence that these processes do not differ between students and other populations. If there is a lack of this evidence, a pilot study comparing a small sample of students against a small sample from the population in question could be carried out to analyze whether the processes differ or not (Lamb and Stem 1980). Student samples may also be adequate in pretesting or in exploratory studies. However, using them for this purpose may also be problematic, for instance, the use of a student sample to select stimuli or product categories for a main study with adults. Students may have a totally different perception of the nature of a stimulus or, say, their involvement with a product category, simply because of the context they are in, and differences between 'youth culture' and how adults stand in life. Calder, Phillips, and Tybout (1981) make the distinction between what they call 'effects application' and 'theory application'. The former aims at statistical generalization of a theory. In

that case, a close match between the research sample and the population of interest is required. Student samples would most often be inappropriate in this case. The latter aims at theory falsification, for which any (preferably homogeneous) sample in the theory domain, and thus also students, is suitable. It goes without saying that advertising research falls into the first category. One may assume that advertising research aims at understanding the responses of a large variety of 'ordinary' customers, and is not (or should not be) primarily interested in falsifying theories for its own sake. In the articles I screened, the latter is also never put forward as a research objective. To add insult to injury, the closing sections of these articles never argue that 'when targeting students, advertisers should…'. Conclusions and recommendations are invariably posited as universally relevant. How cost- and time-saving they may be, be careful in using student samples in advertising research and justify your choice.

In all research with samples of participants, basic rules of good sampling techniques should be observed. Samples should at least be relevant for the topic studied, preferably representative of a population of interest, and random were needed. Sarstedt et al. (2018) observe that most advertising studies either lack information on the sampling method used or use questionable sampling methods that lead to potentially biased or inappropriate samples of participants. They urge researchers to keep the fundamental aspects of sampling in mind to increase the rigor and relevance of their results.

More than half of the articles I screened (both surveys and experimental studies) are carried out online. The obvious reason is that online data collection is fast and cheap, it allows the selection of relevant and representative samples, and interviewer bias is avoided. However, they pose a number of specific problems that should be taken into consideration and solved. Researchers

should realize that they have less control over the conditions in which respondents fill out an online questionnaire, due to the physical distance, lack of personalization and distraction. If, for instance, an experiment that involves mood manipulation is run online, and respondents take a break in between the mood induction and completing the dependent variables, than the data becomes useless as the mood manipulation will not last long enough to affect the dependent variables. Also, multi-tasking (e.g., listening to music or watching TV while completing the questionnaire), the presence of friends or relatives, distracting or annoying background noise, etc., may all render manipulation efforts ineffective (Geuens and De Pelsmacker 2017). Chandler, Mueller and Paolacci (2014) found that participants in online studies admitted to frequently multitask, with 18% of them watching television, 14% listening to music and 6% instant messaging while filling out a questionnaire. Controls should therefore be built in to identify and eventually remove careless or inattentive respondents, such as set minimum and maximum duration times for exposure to stimuli, prevent inattentive responses via instruction sets, 'bogus' instructed response items (for instance, 'if you are completing a questionnaire now, respond with a six for this item'). Respondents who fail these checks better be excluded from further analyses (Geuens and De Pelsmacker 2017). See Kees et al. (2017) for a more in-depth discussion on sampling and data quality.

Novelty Seeking?

There is popular Dutch saying that translates in English as 'act normal, that's already crazy enough', which appears to be very applicable to advertising researchers. Academics often stick to a limited set of pet methodologies, illustrating the principle that, if you only have a hammer, you

see nails everywhere. Of course, it is not surprising, and obviously not problematic, that a limited number of appropriate methods are used in advertising research. Structural Equation modelling is used in 11% of the articles screened, 15% uses PROCESS models, and 16% some other form of regression analysis. The vast majority of the rest of the studies uses an experimental approach and analysis of variance (for a structured overview of research methods used in advertising research, see Chang 2017).

In a previous section, I pointed at the sometimes problematic practice of tossing around correlations in regression-based analyses. However, experimental studies, the 'golden standard' of consumer research, are not without problems either. As Bergkvist, Hjalmarson and Mägi (2016) eloquently put it: they only provide short-run answers to short-run stimuli. Most of the time, the findings are not externally valid in any way, because of artificial research contexts, such as consumer labs and completing online surveys while watching television, because the stimuli used are not realistic, the exposure or consumption context is ignored, and/or effects of advertising are often only relevant if they persist over a longer period of time (longitudinal effects are not typically studied in experimental advertising research). Additionally, researchers also assume that consumers are capable of validly and reliably answering whatever question about anything. I recently came across a study that measured scores on 32 sensorial characteristics of a piece of recycled plastic. In another one, consumers had to score 10 discrete emotions after briefly being exposed to a new label on a bottle of a new brand of beer. On top of that, researchers often get overly excited about small but significant differences between experimental conditions, usually because of overpowered studies. Again, don't get me wrong. Most of these pieces are written by academics with impeccable research skills, but some of them would benefit from taking a step back and ask themselves: does the way I am going to do this really reflect reality?

About 14% of the articles I screened used some kind of 'novel', or let's rather say 'less conventional' method, and the numbers have gone up over (an admittedly short period of) time. Maybe you will be disappointed by the examples that follow because they are, in fact, not so 'novel' at all. But they are something else than the mainstream methods I mentioned in the previous paragraphs. Most of these less conventional methods are related to 'big data' type of studies, such as sentiment analysis, text mining, and online behavioral data. Obviously, many 'big data' researchers, even in advertising, publish their work in other journals than IJA and JA, but it is a pity that often they don't seem to consider advertising journals as a suitable outlet for their work. Other examples of methods used in advertising research that are relatively novel compared to traditional survey-based or experimental studies, are eye tracking, neurophysiological measurements, virtual reality techniques, face analysis software, response latency measures, online focus groups, and methods that try to enhance real-life contexts, such as living room studies. I am not arguing that all advertising researchers should all of a sudden be obsessed with 'unconventional' research methods, but it would help if researchers at least consider them to make their work more interesting or realistic. Remember: advertising takes place in the wild, not in a consumer lab.

It ain't over till the fat lady sings

This proverb refers to opera sopranos, who were traditionally overweight (nowadays they come in all shapes and sizes). It means that one should not assume to know the outcome of a situation or event while it is still in progress, because the context can change (https://knowyourphrase.com/aint-over-until-the-fat-lady-sings). This is particularly relevant for advertising. Nothing has happened until the sale is made. The vast majority of outcomes studied

in advertising research do not explore what people actually do, but only what they think, feel and/or intend to do, and readily assume that this is predictive of real behavior. This is so common in advertising research that most authors not even bother to mention it as a limitation, and when they do, at best they only vaguely refer to studies that show a strong correlation between people's thoughts, feelings, attitudes and intentions on the one hand, and real behavior on the other. However, attitudes or intentions are do not necessarily lead to behavior, and 'strong' is relative: a correlation of 0.5 between the two means that part of the variance in behavior is due to other factors than intentions, such as habits and context, for instance. Even the correlation between stated and real behavior is often not higher than 0.70 (see for examples in a driving behavior context: Cauberghe et al. 2009). Indeed, the notorious intention-behavior gap is a major research topic (see, for instance, Carrington, Neville, and Whitwell 2014). Knowing this, why are researchers only interested in what consumers think, feel and intend to do? The answer is simple: what else are you going to measure in a survey or an experiment, unless some academic charity is going to donate you a nice amount of money to additionally study real behavior in a real context?

A related problem is the issue of construct measurement in advertising studies. There are substantial inconsistencies between different scales that are supposed to measure the same construct (e.g., great diversity in the items included in measures), antecedents and consequences of the particular construct are included in the measures of the construct, and the number of items considerably varies between different scales for the same construct (Bergkvist and Langner 2017). This compromises the validity and comparability of advertising research and hampers the accumulation of knowledge and progress in advertising research.

About 10% of the screened articles use some kind of 'real' data. Sometimes these data do not actually measure actual consumer behavior, but the result of this behavior, for instance advertising

budgets spent and resulting sales. On the other hand, there is increasingly more behavioral data available, especially online, that are very relevant for advertising researchers and advertisers. The majority of the research using real behavior data are big data studies in which, for instance, liking, sharing, comments and clicks are measured and linked to advertising messages on social media platforms. One may question the predictive value of this 'consumer engagement' data for actual buying behavior, but at least these actions actually happened in a real context. Occasionally, also experiments measure real behavior (such as choosing a real product after exposure to advertising stimuli), but those studies suffer from the shortcomings discussed in the previous section. Increasingly, top journals demand also real behavior studies as part of a research project. Advertising researchers should include them in their studies too. Read Malthouse and Li (2017) on how to use behavioral data in advertising research.

To measure is to know, but to talk is to understand

Only a mere 5% of the articles I screened report qualitative studies, mostly as part of a pre-test or scale development, or when studying industry practices (for instance, interviews with advertising professionals). Do advertising researchers not like qualitative studies? Or editors and reviewers? There are other disciplines in which qualitative research is highly regarded and widely published. Maybe it has to do with the type of people that do advertising research and how they are trained, or we do not feel comfortable about the findings of qualitative studies. They are not full of little numbers we can torture until they confess. Or as one of my colleagues puts it: I don't like qualitative studies. In experimental studies you can always find a number here or there you can build a story upon. Most probably, authors refrain from submitting qualitative or mixed methods

studies because it is practically impossible to get such studies published, mainly due to 'ideological' principles of editors and reviewers.

I have always found qualitative studies interesting. You often learn more from observing a focus group discussion than from a carefully crafted experiment or a survey. A least, if well done, qualitative research can lead to original research ideas or deepen our understanding of how real people feel, think and behave. In times of real-time programmatic advertising and A/B testing of online ads, and of online number crunching, advertisers are often shooting from the hip, without much understanding why things are going on. In this context, qualitative research can be very important to acquire a deeper understanding of consumer responses and behavior. For an insightful overview of the role and importance of qualitative advertising research, see Belk (2017).

Literature reviews and research agendas are cool

About 15% of the articles I screened are literature reviews, meta-analyses or conceptual papers that usually also present future research agendas. They are important. Of course, there needs to be a sufficient amount of original research articles before any meaningful literature analysis can be conducted and further research proposals can be developed. But once in a while, it is necessary to take a step back and focus on the integration of knowledge rather than producing a plethora of incremental contributions to it. A well-executed literature review or conceptual piece, including an inspiring and detailed research agenda, can be very useful for researchers. It enables them to quickly and efficiently get a grip on the state-of-the-art of past and current research about a topic, and it is a great resource for scholars who start doing research in a domain they are not yet familiar

with (for instance, Ph.D. students). Additionally, it provides inspiration to identify important gaps in current knowledge and build your own research agenda.

The academic bubble

Could it be that editors, reviewers and readers are too much locked up in their academic bubble? Authors choose research approaches and ways to report their results in function of their perception of what reviewers and editors will like: a manuscript should tell a smooth, consistent story, with no contradictory findings, and preferably all $p < 0.05$. It is obvious that such papers have a greater chance of being published than work that does not meet the expectations of editors and reviewers.

Fiedler and Schwarz (2015) tested the prevalence of ten questionable practices in psychological research. Most of them are also relevant (and probably prevalent) in advertising research. Researchers report doing the following five practices in between a stunning 30 and 50% of the time: claiming to have predicted an unexpected outcome, selectively reporting studies regarding a specific finding that 'worked', deciding whether to exclude data after looking at the impact of doing so regarding a specific finding, failing to report all dependent measures that are relevant for a finding, and collecting more data after seeing whether results were significant in order to render non-significant results significant. Authors don't do that for fun, but because they expect that this is going to increase their chances of being published (and they are probably right).

Most of these practices have to do with two phenomena: p-hacking and HARKing. P-hacking refers to manipulating reported significance levels by rounding them off inappropriately (for instance, reporting a p-value of .054 as .05), remove 'outliers' or collect additional data to push p-

values below the sacrosanct 0.05 threshold, or carefully select dependent and control variables to report results that meet desired significance thresholds. HARKing means Hypothesizing After the Results are Known (Kerr 1998; Bergkvist 2020).

Removing outliers is, in itself, not a questionable research practice, provided it is done appropriately and reported in full transparency. In some cases, outliers can provide very relevant insights, for instance, in studies analyzing how people generate creative ideas, researchers will find a very small amount of people who generate an enormous number of ideas. These 'outliers' are very relevant. However, selectively or arbitrarily removing data to 'pimp' a p-value or improve results in general, without even mentioning it, is condemnable.

These practices determine how science is conducted and which science gets published. Is that what we want? Why don't we all (editors and reviewers) become a bit more tolerant about not-so-smooth stories that contain contradictory results and in an honest way report findings that are not perfectly in line with the p-value axioma?

One reason why so many submitted or published advertising research has no meaningful contribution to practice, is because it pays off to stay mainstream, keep on doing the same things in the same way as in previous research, and not overly worry about the practical usefulness of the findings for advertising. In their citation analysis of JA, IJA and JAR articles, Chan et al. (2017) expected that articles that are related directly to advertising practice, effects and content would get more citations than topics which focus on peripheral areas to advertising. The contrary appeared to be the case. Probably, editors and researchers are aware of this, and may therefore develop the attitude that 'peripheral' advertising research is better for their careers and the reputation of their journals than practically relevant research. This is one of the consequences of the 'publish or

perish' pressure in academia. Most universities only focus on the number of papers and the journals these papers have been published in when hiring and evaluating scholars. Practical relevance of research (and teaching) is most often not considered. However, we should not only worry about academic impact, but also (and maybe even more) about the impact of what we do on the real world.

Call to action

As mentioned in the introduction, it is not my intention to nail the whole advertising research community to the pillory. Many of us are excellent researchers. But at the same time, probably all of us have fallen into some (or most) of the traps I discussed, including myself: guilty as (not) charged. There are several measures that can be taken to address these issues, and editors have an crucial gatekeeper role to fulfill.

First of all, editors could take the initiative to initiate special issues of their journals that focus upon topics that deserve more attention, broaden the scope of advertising research and focus upon topics that are important for contemporary advertising practice. Fortunately, they already do. In recent years, IJA and JA hosted special issues on, amongst others, new trends in digital and social media advertising, digital engagement with advertising, native and covert advertising, reinquiries (replication) in advertising research, literature reviews, leveraged marketing communication, social and environmental issues in advertising, big data in advertising research, artificial intelligence in advertising, future direction in advertising research, and advertising in Asia and in Latin America.

Second, editors have an important lever to impose or discourage certain practices: author guidelines. Instructions for authors could be adapted to emphasize desirable and undesirable features of submitted manuscripts (instead of bothering authors with instructions about title formats, word count for abstracts and detailed instructions on how to upload tables and figures, to name a few). For instance, they could clearly state that submissions without a specific, convincing and actionable managerial implications section or that are overly based on student samples will be in trouble. They could explicitly emphasize that literature reviews, conceptual articles and research agenda contributions, replication studies and qualitative studies are welcome, and that novel methodological approaches and (also) measuring real behavior are a plus, without falling into the trap of ignoring sound theory building and application. Editors then also have the duty to educate their reviewers (all of us, that is) to take these guidelines into account. And obviously, they have to put their money where their mouth is, and steer the review process and take editorial decisions accordingly.

Editors could organize a small revolution. Bergkvist (2020) suggest preregistration of advertising studies to take the angle out of HARKing and p-hacking. Preregistration means that authors develop a conceptual framework and hypotheses and report how they will collect data, prior to actually carrying out the study. This pre-paper is then stored on a preregistration platform, and eventually, upon submission, sent to reviewers along with the full paper. This is an attractive option. In that way, hypothesis testing is based upon prediction rather than ‘postdiction’ and, consequently, the HARKing and pimping p-values practices are discouraged (unless authors write the pre-paper after having collected and analyzed the data, that is, but that could be prevented by data stamps on the pre-paper and on data collection documents). As some journals in other disciplines do, one could even go a few steps further. Authors could be asked to also develop a

preliminary ‘managerial contribution’ section in their pre-article. Pre-articles could also be formally reviewed by the journal. Once accepted, the authors can carry on with their work. After completion of the full paper, they could submit it along with their accepted pre-paper. Preferably the same reviewers as the ones who looked at the pre-paper could review the full paper. They would then not look anymore or question the conceptual development and the hypotheses, and papers would be acceptable for publication, even if hypotheses are not confirmed or not all effects are significant at conventional levels. The existing preregistry deposits time stamps preregistrations, so these are hard to antedate after the fact. However, it is possible to cheat with when the data were collected (e.g., by changing or removing all dates in the datafile). However, this procedure would reassure authors that the acceptability of their paper does not depend on what they choose to report and how, and therefore they could be less tempted to cheat.

Finally, what we may need is a cultural change in the way we hire and evaluate researchers and professors. Consistent with the ‘publish or perish’ obsession, most universities only focus on output of papers and the journals these papers have been published in. Practical relevance of research is most often not considered. To change this mindset would be very difficult, since it is firmly engrained in the academic mindset and often considered self-evident. Maybe a first step could be that academics who are involved in hiring and evaluating faculty take managerial relevance of a researchers’ work into account. What might also help is compulsory internships in an ad or media agency for advertising researchers and Ph.D. students.

Conclusion

There is a lot of room for improvement in academic advertising research. Lack of realism and

managerial relevance, a focus on theoretical and methodological sophistication at the expense of practical applicability of research outcomes, incremental contributions instead of integration of knowledge and dissemination thereof in the world of advertising practice, a lack of interest in trying to convince the professional advertising sector of the relevance of what we do but, instead, remain locked up in the moral high ground of academia, remain existential challenges to the advertising research community. If we want to stay or become relevant to the society that pays our salaries, we will have to do something about it.

Fortunately, let's not forget that a lot of good things already happen in advertising research: many papers do report relevant and actionable suggestions for practitioners, researchers increasingly use novel methods and report behavioral data, and editors do make efforts to encourage research on contemporary and important advertising issues and practices. Let's keep up the good work!

Declaration of interests: No conflicts of interests.

References

Belk, Russell W. (2017) Qualitative research in advertising, Journal of Advertising 46:1, 36-47.

Bergkvist, Lars (2020) Preregistration as a way to limit questionable research practice in advertising research, International Journal of Advertising, DOI:10.1080/02650487.2020.1753441.

Bergkvist, Lars, Hanna Hjalmarson, & Anne W. Mägi (2016) A new model of how celebrity endorsements work: attitude toward the endorsement as a mediator of celebrity source and endorsement effects, International Journal of Advertising 35:1, 171-184.

Bergkvist, Lars & Tobias Langner (2017) Construct Measurement in Advertising Research, Journal of Advertising 46:1, 129-140.

Calder, Bobby J., Lynn W. Phillips, & Alice M. Tybout (1981) Designing research for application, Journal of consumer research 8:2, 197-207.

Carrington, Michal J., Benjamin A. Neville, & Gregory J. Whitwell (2014) Lost in translation: Exploring the ethical consumer intention–behavior gap, Journal of Business Research 67:1, 2759-2767.

Cauberghe, Verolien, Patrick De Pelsmacker, Wim Janssens, & N. Dens (2009) Fear, threat and efficacy in threat appeals: Message involvement as a key mediator to message acceptance, Accident Analysis & Prevention 41:2, 276-285.

Chan, Terri H., Caleb H. Tse, & Kineta Hung (2017) Productivity and impact in advertising research since the millennium: a profiling and investigation of drivers of impact, International Journal of Advertising 36: 1, 11-37.

Chandler, Jesse, Pam Mueller, & Gabriele Paolacci (2014) Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers, Behavior Research Methods 46: 1, 112-130.

Chang, Chingching (2017) Methodological issues in advertising research: Current status, shifts, and trends, Journal of Advertising 46:1, 2-20.

Erfgen, Carsten, Sebastian Zenker, & Henrik Sattler (2015) The vampire effect: When do celebrity endorsers harm brand recall?, International Journal of Research in Marketing 32:2, 155-163.

Fiedler, Klaus & Norbert Schwarz (2015) Questionable Research Practices Revisited, Social Psychological and Personality Science 7:1, 45-52.

Fleck, Nathalie &Virginie Maille (2010) Thirty years of conflicting studies on the influence of congruence as perceived by the consumer: Overview, limitations and avenues for research, Recherche Et Applications En Marketing (English Edition) 25:4, 69-92.

Fuchs, Sebastian &Marko Sarstedt (2010) Is there a tacit acceptance of student samples in marketing and management research? International Journal of Data Analysis Techniques and Strategies, 2:1, 62-72.

Geuens, Maggie & Patrick De Pelsmacker (2017) Planning and conducting experimental advertising research and questionnaire design, Journal of Advertising 46:1, 83-100.

Henrich, Joseph, Steven J. Heine, & Ara Norenzayan (2010) The weirdest people in the world?, Behavioral and Brain Sciences 33:2-3, 61-83.

https://knowyourphrase.com/aint-over-until-the-fat-lady-sings (Accessed 26 May 2020).

https://www.britannica.com/topic/Occams-razor) (Accessed 26 May 2020).

Jones, William L. & Brenda S. Sonner (2001) Just Say No to Traditional Student Samples, Journal of Advertising Research 41:5, 63-71.

Kees, Jeremy, Christopher Berry, Scot Burton, & Kim Sheehan (2017) An Analysis of Data Quality: Professional Panels, Student Subject Pools, and Amazon's Mechanical Turk, Journal of Advertising 46:1, 141-155.

Kerr, Norbert L. (1998) HARKing: Hypothesizing after the results are known, Personality and Social Psychology Review 2:3, 196-217.

Lamb Jr, Charles W., & Donald E. Stem Jr. (1980) An evaluation of students as surrogates in marketing studies, Advances in Consumer Research 7:1.

Malthouse, Edward C. & Hairong Li (2017) Opportunities for and pitfalls of using big data in advertising research, Journal of Advertising 46:2, 227-235.

Royne, Marla. B. (2016) Research and publishing in the journal of advertising: making theory relevant, Journal of Advertising 45:2, 269-273.

Sarstedt, Marko, Paul Bengart, Abdel Monim Shaltoni, & Sebastian Lehmann (2018) The use of sampling methods in advertising research: A gap between theory and practice, International Journal of Advertising 37:4, 650-663.

Schultz, Don (2016) The future of advertising or whatever we're going to call it, Journal of Advertising 45:3, 276-285.

Taylor, Charles R. (2020) Common mistakes made in submissions to leading advertising journals, International Journal of Advertising 39:3, 329-333.

Völckner, Franziska, & Henrik Sattler (2006) Drivers of brand extension success, Journal of Marketing 70:2, 18-34.

Word count: 6,750