Dear Author(s):

Your submission to the Journal of Marketing Research, JMR-12-0061.R1, entitled "The Limits of Attraction," has been reviewed. The comments of the reviewer(s) are included at the bottom of this letter.

The revised version of your paper was read by the same two reviewers and AE who were involved in the last round, and, being new to the paper, I read both this new version and the earlier one as well as both sets of reviewer comments. Let me cut to chase by saying that the call on this one is not easy, as the reactions by the review team could not be more varied, ranging from conditional accept (R2) to reject (AE). Here is the landscape as I see it. I think everyone would agree that no other “effects” paper in the BDT area has attracted more interest, controversy, and used up more lab subjects than Joel and Chris’ original 1982 attraction effect paper. Unlike the experimental results leading to Prospect Theory, which one could view it as a not-totally unfriendly refinement of utility theory, the attraction effect was a result that, if true, threw a fundamental monkey’s wrench in random utility theory, which was (and still is) the analytic backbone of modern IO economics and a large chunk of the marketing research field. As a result—and I remember this quite vividly—it was a result that many suspected was a lab parlor trick that would not replicate when put under close inspection. The most common argument was that since standard RUMs did quite well (thank you) in predicting choices from real-world sets, it must be something about the artificiality of the lab choice sets that was doing it, where, for example, “quality” was defined by an ambiguous numeric score. I am pretty old, and recall talking to Allan Shocker at a conference around 1984 or so telling me that he didn’t buy it, and was poised to write a paper showing that all one needed to do was swap out the abstract options for “realistic” ones and sure they effect would do away. Problem was when they did that study it did NOT go away (Ratneshwar, Shocker, and Stewart ’87). My sense of the current status of the effect is that most accept it as very real but also very limited, a result that plays a role in choice theory akin to empirical demonstrations of bending light in physics: it is helpful in showing the Newtonian laws (here, RUT) do not explain the universe, but the domain is sufficiently limited that for most practical applications standard theory works fine (hence Sawtooth is not ripping clients off).

**I agree with your characterization of the effect, though not your characterization of the popular perception of the effect. Understandably, when best-selling books within our field speak of these massive effects in realistic contexts, prior suggestions regarding the breadth and depth of the effect are prone to be substantially exaggerated.**

Against this background, like the AE, I found this paper to be a puzzle. It’s quite well done, and written with a populist, revisionist, tone that underlies many recent attempts to reconsider classic social psych results (like old-age priming). But unlike those efforts the reexamination here is not a new one; it reads more as a throw-back that resurrects many of the debates that occurred in the 80s on the robustness of the attraction effect, but without really acknowledging them, or explicitly building on them. What’s also odd is that the AE raised this same point in the last round, but you did nothing with it, preferring to plow on as if you were the first to seriously challenge the empirical status if the attraction effect, and the first to conclusively debunk it.

**I don’t think this is accurate. I believe there are five papers that can be considered serious attempts to examine the boundary conditions of the effect: Kivetz, Netzer, and Srinivasan, 2004; Ratneshwar, Shocker and Stewart, 1987; Sen, 1998; Simonson and Tversky, 1992 and Trueblood, 2013. RSS antedates the others, but the other papers are just as much a test of the boundary conditions of the phenomenon – indeed, the stimuli used by RSS represent a comparatively modest departure from the typical studies with highly stylized stimuli. Those other studies claim that the attraction effects persist in contexts departing *further* from the historical benchmark than RSS itself.**

**In this sense, I find it odd to dismiss the evidentiary value of our paper by saying that “RSS already did this.” Consider the following analogy. Study 1 shows that mixing in some breast milk with formula reduces autism. Studies 2-5 show that breastfeeding has no effect on autism. Study 6 shows that breastfeeding eliminates autism. In the context of studies 2-5, study 6 retains evidentiary value, even if study 1 had earlier found results that are qualitatively consistent with (though weaker than) study 6. Moreover, to the extent that design flaws in Study 1 limit its evidentiary value, study 6 is even more important. We did not dwell upon the flaws of study 1, because we think it coincidentally arrives at a qualitatively correct conclusion, though I don’t think it goes far enough, nor does it emphasize what we see as the essential distinction. More on this later.**

**If you may permit a grandiose analogy, I feel like the AE repeatedly upholding RSS as grounds for rejecting this paper is a bit like rejecting Prospect Theory on the grounds that Markowitz already mentioned something about reference dependence in a prior paper. Surely, there is room for more than one paper on a topic. We succinctly summarize a massive amount of data using a broad array of stimuli, which represent a much more aggressive departure from the status quo than RSS.**

Now, what you do here is much more comprehensive than what anyone has done before, but rather than use this impressive arsenal of work to offer a more nuanced (or clarified) view of when the attraction effect works and why, you opt for the more blunt—and less helpful conclusion that it simply is “not true”.

**Though inaccurate, that would not be an altogether inappropriate pragmatic takeaway. That said, this is not the conclusion we draw or report. With stimuli involving 2X2 numeric matrices, we report significant effects in 4 of 5 cases. Though the distinction we emphasize (perceptual vs. numeric) tends to correlate with “realism,” it is distinct. For instance, unlike most choice options, gambles can be represented with perfect *realism* using only numbers. We find strong attraction effects here. The effect is eliminated when one of the dimensions (probability) is represented perceptually (as the visual area of a probability wheel). However, in this case, at least, we don’t believe that is because the stimuli are more “realistic” – indeed, I think they are less so, as the probability wheel representation is rarer in the real world than the numbers themselves. Of course, in most of the remaining stimuli, perceptual representations correspond with greater realism.**

As point of comparison, have a look at what Scheibehenne, et al. (JCR 2010) did with their reconsideration Iyengar’s choice overload effect, which also does not always replicate. What they did was rather than just saying, “it’s not true”, they tried to reconcile the work that had been done via a meta analysis, pointing out that the strength of her effect was extremely rare. I thought you would be doing something like that here---trying to use the lab studies to resolve just what it is about some set ups that allow the effect to be observed and what one needs to make it go away.

**We will look at that paper as a touchstone for some things we might do differently. I am familiar with it, though have only skimmed it. Uri Simonsohn has a nice analysis of the studies reported there, suggesting that those showing the celebrated effect are p-hacked, whereas those showing the more expected effect are not. I might conclude from this secondary analysis that there is no effect to decompose; just file drawers and publication bias.**

**I am less cynical about the attraction effect, since we *do* find the effect using non-bizarre stimuli, such as gambles. But the effect is certainly much narrower than the way it is typically characterized. Certainly, clearer specification of boundary conditions is something toward which to strive. I think this paper makes some progress, but leaves room for more.**

“More realistic” is clearly too blunt (e.g., Ratneshwar, Shocker, and Stewart ’87). For example, one of the long-standing assumptions is that for the attraction effect to hold there must be ambiguity or uncertainty about the meaning of (or preference for) at least one of the attributes; it doesn’t work for money, nor would it work for vivid, familiar, stimuli like preferences for views from apartments.

**I disagree that that is a widely held assumption, and disagree that the typical characterizations of the effects suggest the boundary condition we posit. Nor do I think it perfectly maps onto the data. Upon receipt of the data, the AE and R1 offer strident post-dictions that the results were predictable, but many of the stimuli were constructed in response to asking champions of the effect to design stimuli that would show it, with the constraint that no more than one dimension could be numerically specified. Many candidate stimuli were offered up, though we’ve not yet found one that yields an effect. In some cases, even *these* people later find reasons why these stimuli weren’t right for testing the effect urged us to run a 28th study with their candidate stimulus du jour. Moreover, even if we cede n examples as being, for some reason or reason(s), inappropriate for testing the effect, the remaining 27-n studies still require an account.**

You might have a look at Paulo Natenzon’s paper, “random choice and learning” (working paper Wash U) that lays out a mechanism by which an otherwise rational decision maker who is uncertain about their preferences will display choices consistent with an attraction effect—a boundary condition that might be evident in your data.

**Thanks. We will have a look.**

Likewise, as the AE suggests (also R1), it is important to wear the hat of someone on the other side, someone who believes the many studies that seem to show the effect. Are you sure that, in fact, that people actually perceive the dominance in all these studies?

**Yes. I think there is no other sensible account for the overwhelming preference of the target over the decoy.**

So here is where I think we are, which is basically where we were last time. There a couple of reasons that I would ideally like to see some version of this paper eventually make it in print. There is a widespread sense that we should be encouraging more replication work in our field, and this paper does that, and does it reasonably well. Publishing it would send a nice signal. Then there are the studies themselves, which represents an impressive pile of empirical evidence on one of our field’s most well-known effects. It ought to be in print. All of that said, I agree with both reviewers and the AE that in its current form the paper does not appropriately acknowledge the long history of scholarship that has tried to do the same thing. One might argue that the endeavor is novel because there has been only one major published paper that has tried to look at the effect of using meaningful stimuli, but think about it: once the RSS paper was published, it would have been hard for anyone else to publish the same thing again (it’s hard enough getting a replication study published, much less a replication of a replication). You are not the first to do this, not the first to wonder why we see the effect in the lab with numerical quality scores, but choice models do not seem to suffer from it when predicting real-world choices of toasters.

**Well, I think the autism analogy I drew earlier is relevant here. I don’t think our paper quite settles the matter either, though it makes a substantial advance.**

**Regarding RSS, I think it needs to be considered in the context of the other papers that have examined boundary conditions (around six such papers including ours). I’ll reprise the questions I raised in my email to you**: (**1) Is that paper enough?  (For someone interested in this phenomenon and where it works and doesn’t, does this paper provide all the answers?)   (2) Are *those* results replicable?  (3) Are the designs there appropriate? (4) Is the authors’ account of their results accurate? (5) Are the manipulations conducted there “essentially” the same as those we use. One can, for instance, distinguish between *supplementing* numbers with verbal descriptions (e.g. Ratneshwar, Shocker and Stewart, 1987), *replacing* numbers with verbal descriptions (e.g., Sen, 1998), and replacing numbers with perceptual representations (Frederick, Lee, and Baskin, 20XX).**

**We’ve actually conducted some direct and conceptual attempts to replicate RSS. I will have to excavate the details, but the upshot is that they don’t look especially similar. I’m not convinced that *supplementing* numbers with verbal descriptions typically has the same effects as the sorts of manipulations we use. For reasons I can elaborate upon later, I don’t believe their effects are, in fact, caused by enhanced meaningfulness; I don’t think that verbal descriptions appended to the numeric summaries actually enhances meaningfulness very much in that context. Respondents get it; these are three orange juices that can be ranked in quality and price; the higher priced ones with higher ratings presumably taste better and the supplementary verbal descriptions confirm this – though they have other confounding effects, including suggesting that a property of the cheapest orange juice is shared by the intermediate one.**

**We avoided extensive reference to this paper because I think it is the worst of the five that precede ours, because it uses an unnatural manipulation that essentially preserves the 2X2 numeric matrix from which we wanted to depart, because it uses a manipulation that we believe would be largely ineffectual save for an artifact caused by an inappropriate design, and because the results don’t really replicate anyway. Furthermore, as noted earlier, it seemed odd to spend time critiquing this study specifically, when subsequent studies using *stronger* manipulations (i.e., more realistic stimuli) appear to show that the effect was *not* attenuated. Of course, we’d be happy to more explicitly acknowledge the temporal priority of this work, and to discuss and critique it at length. The later studies were more pertinent to our central interest, because they represented more substantive departures from the typical design, and, thus, a more ambitious exploration of the envelope of the effect.**

To be a contribution to the literature in this area you need to build a more explicit case for how the findings BUILD on what has been done before. It’s not enough simply to say, “it doesn’t work for window views”. Why doesn’t it? Or, alternatively, what is it about numerical quality scores that causes it to work? Go back to the physics example. As I mentioned why people think the AE is important is not that it discredits standard modeling practice, but rather that gives an insight into a nuanced feature of the psychology of choice that standard field data would be too coarse to pick up. It seems that with the vast amount of data you have here you ought to be able to tell us much more about what the original AE designs were saying about choice psychology than what we had before.

**We speculate about the reasons, but fail to isolate the precise psychological mechanisms. I agree the paper falls short of our ideals in this respect – perhaps due to lack of imagination, though certainly not from lack of effort or lack of familiarity with similar work. But I think the paper goes a long way toward establishing the contours of the boundary, and that it will foster further work to help settle their exact location.**

Hence, while there has not been great convergence in this round, I would nevertheless like you to try again to write a paper that addresses the concerns raised by myself and the review team. Ideally it’s a paper we would like to publish, but you have to do your paper by writing a work that truly builds on prior efforts, not works in parallel to it.

**I think it *does* build on prior results (though sometimes on the location where they’ve been razed). This seems like a good departure point for a conversation (though I’ll be very sleepy).**

**Regards, Shane.**

As before, your revision must be submitted through Manuscript Central within one year and should include no more than 4-5 pages of notes characterizing the changes you made in the manuscript and responding to the specific reviewer comments. Exceed 4-5 pages only if you feel it really is necessary to include more detail.

Please remember that your revision must adhere to our 50 page limit (inclusive of title, abstract, text, references, tables, figures, footnotes, and print appendixes). Information that is not central to the understanding of your article should be moved to a Web appendix.

We look forward to reading the paper.

Sincerely,

Prof. Robert Meyer

Editor

Journal of Marketing Research

[meyerrjmr@wharton.upenn.edu](mailto:meyerrjmr@wharton.upenn.edu)

Associate Editor's Comments to Author:

Associate Editor

Comments to the Author:

AE report for Ms 12-0061-R1

On the last round, I made the observation (as also pointed out by reviewer 1) that your paper sounds very similar to the Ratneshwar et al. 1987 paper. Yet, you disagree with me. I have no problem with a disagreement if you can give me a convincing argument. Yet, I see none here. In fact, as far as I can tell, you now don’t mention the Ratneshwar et al. 1987 paper anywhere in the text.

So, let me go back to that original paper and quote from it, and let’s see if it sounds similar to what you are doing:

a. “The attraction effect refers to an inferior product's ability to increase the attractiveness of another alternative when the inferior product is added to a choice set. This article examines potential explanations for the attraction effect and its boundary conditions. The article reports several empirical investigations and suggests that the attraction effect may be moderated by such variables as stimulus meaningfulness and familiarity with the product category. The implications are relevant to research on context effects in consumer choice.” (Abstract)

b. “H: Elaboration that leads to greater meaningfulness in the stimulus brand descriptions results in reduced attraction effects on choice. Given meaningful stimulus descriptions and a high level of product familiarity, the attraction effect will not be significant. (p. 524)

c. “The attraction effect was replicated in those conditions of the study where the descriptions were ambiguous and lacked sufficient meaning, but the effect size was substantially diminished when the descriptions were elaborated and made more meaningful” (p. 531)

You quote Ratneshwar et al. in the reviewer notes (but not in the paper itself), who say that:

“Future researchers examining context effects…should create more realistic situations…” but the reason they say this it NOT because no one has done this, but because their own results show this to be important. I am just at a loss as to how you can utterly ignore this research, when you show that the attraction effect is reduced in more realistic situations and Ratneshwar show that the attraction effect is reduced in more realistic situations. How is that not the same idea?

I understand that you position the paper differently from Ratneshwar et al. Ratneshwar et al. positioned their idea as a boundary condition/moderator, whereas you position the paper around the idea that discussion of the attraction effect should be removed from consumer behavior courses. I guess, if anything, you have confirmed that Ratneshwar et al. were correct (long ago), and maybe your paper could be summarized as, “All of you who replicated the attraction effect with numeric representation, and wanted to show that your results are externally valid (rather than making a theoretical point), should have listened to Ratneshwar et al. and used more realistic descriptions.”

I also had asked you to provide more details on your stimuli, and I am happy to see that you did this. But, I had also asked you if respondents could detect dominance relationships with the more realistic stimuli (as also noted by reviewer 1). You said that, sure, you think respondents can detect these relationships, even with more realistic stimuli. But, what evidence can you offer for this? Again, I can see that a dominance relationship is pretty apparent with numerical stimuli (I can readily see, for example, that option A gives me a better price and is better on an attribute that I care about), but I’m just not sure this is true for the stimuli you used. For example, let’s look at p. 32 of your paper, where you tested Jalapeno flavored popcorn. Do people really think that option B here dominates Option C? It seems to me that consumers would simply think that option C is not of interest to them, as it is targeted after someone who likes spicy popcorn, but NOT that Option B dominates Option C. And, what about p. 31 – do people really pick up on the fact that Speed 2 is dominated by Speed 1? And, what about p. 34, where you look at Duck Fart water? You are assuming here that the consumer thinks Option B dominates Option C, since they are both spring waters, while Option A is regular water. But, again, I just have trouble believing that consumers would perceive this relationship. And, you do get a repulsion effect here, meaning at least consumers do naturally pick up on disgust.

What I am trying to say here is that, don’t we need to assume that consumers first perceive something in order for it to influence them? For example, what if you found that a change in price had no effect on demand? Does that mean that basic economic principles are flawed, or could it be that in this situation consumers did not notice the change in price in the first place? In other words, it is certainly true that in many experiments, we make certain relationships or tradeoffs very apparent, and maybe far more apparent than they would be in the real world, where many attributes are at play. But, we are doing this to get people to at least focus on a certain tradeoff first, and then if they do see this tradeoff, we can measure how they react.

So, in the end, you chose to “double down” as you say in the note to the Editor, but I am also choosing to “double down” on my original suggestions because you have not convinced me otherwise.

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author

The revised paper presents a great deal of new data. It is an interesting paper that contains some intriguing results. However, the reader is left to wonder what the main contribution is and what underlies the findings. Also, the main conclusion is, in my view, overstated. Overall, I think that a rewritten, reorganized paper, possibly with additional data (while deleting some of the current studies), can make a nice contribution.

1. The authors focus on the weakest demonstrations of the attraction/asymmetric dominance (hereafter AD) effect, but the stated conclusions go well beyond these sets. That is, the stated conclusion is that the AD effect is limited to stylized numerical stimuli and does not exist with more realistic stimuli (the specific nature of the difference between the badand good stimuli remains unclear – are the better stimuli more realistic, rich, sensory, or vivid?). Although I share the authors’ concerns about the numerical/stylized sets, the conclusion that the AD effect does not happen in reality is greatly overstated. I assume that the authors do not deny the reality that people assess the value or attractiveness of one option relative to another, as opposed to relative to some absolute standard. For example, when evaluating a camera that comes with valuable accessories and costs just $10 more than the same camera without these accessories, there is little doubt that many consumers would find the former more attractive than they would have in the absence of the relatively inferior reference point. Such effects, which are included under the AD effect, happen all the time (see, e.g., what happens on Amazon, such as where a new item costs the same or just slightly more than a used item). Therefore, I disagree that the AD effect is limited to cases where attributes are rated on 100-point scales. Also, there are numerous sets in everyday life where options are evaluated in terms of a quantitative dimension (e.g., price or, less often, magnification power) and another dimension such as a feature difference. In these cases, the AD effect is very real. I think the authors should be careful in defining what they claim and show.

2. One of the nice aspects of this paper is the many neat choice sets that the authors used (indeed, the web appendix will represent a significant contribution in its own right). One thing I found both interesting and frustrating is the discussion and evidence pertaining to the repulsion effect. Prior research has attempted to demonstrate such an effect (e.g., unpublished work done by Frederick and by Simonson and Tversky). However, repulsion and attraction tend to cancel each other, which makes it challenging to demonstrate a robust repulsion effect. That is, while some subjects “fall” for the AD effect (i.e., choose the option that looks relatively attractive), others “fall” for the repulsion effect (i.e., get so disgusted with the decoy that they avoid everything that looks even remotely similar to it). Looking at the results reported in this paper, one has to wonder whether the realistic stimuli (e.g., a moldy orange) are better at eliciting repulsion than numerical stimuli. If true (and testable), this can represent a contribution. Although the authors mention repulsion in a couple of places, the reader is left wondering what, if any, role repulsion plays in the reported results.

3. One thing the authors can and should do, which will go a long toward understanding why their stimuli do not produce AD effect, is to find out how the relations between the dominating and dominated option (in the more realistic sets) are perceived. If I understand the proposed theory, it appears to suggest that, unlike the stylized stimuli, there is no spontaneous perception of relative inferiority. This account needs to be supported with proper tests.

4. Relatedly, the authors need a more coherent theory that goes beyond saying that the authors see no reason why AD perceptions will emerge with their more realistic stimuli. My guess is that the explanation may be even simpler and more general – the AD effect is weaker with any source of “noise” that makes it less likely that subjects/consumers will notice the set configuration (and in particular the AD relation). For example, with more than two attributes and/or more than three options, the noise tends to make the effect weaker.

5. Some of the studies in the current paper are noteworthy. I particularly like the gambles studies; the TV studies are also nice, though looking on the copy we received, I could not detect any AD in that set (on my screen, the decoy looked good). I also like the study that shows the effect of the rating task order, though this finding might have been shown before (not sure). If it is new, it deserves more emphasis, especially if it can be better integrated in the paper.

6. As an aside, the Ratneshwar et al. (1987) paper deserves more attention because it tried to advance a similar argument (though it was weaker and less persuasive, not to mention that despite their intent to show that the effect would not replicate, it was still observed). Like some of the studies in this paper, Ratneshwar et al. did not confirm the equivalence between their verbal description sets and the numerical sets (and a subsequent test presented at the 1988 ACR showed that the two versions were not equivalent). On a minor note, one thing that the Ratneshwar et al. study did is to confirm that the AD effect replicates with the stylized version. In the current paper, some of the reported studies did not do that.

7. As you can tell from my somewhat disorganized comments, there are many things I like about this paper, but I wish the authors could organize it in a more coherent fashion, strengthen the theory and related evidence, delete or move to the web appendix some of the redundant studies, avoid overstating their conclusion, and clarify the role of repulsion (is that the real driver of the failure to replicate the AD effect?). With these changes, I think this can be an important contribution.

Reviewer: 2

Comments to the Author

This paper continues to improve. However, its tone is needlessly strident, implying that some of the work on asymmetric dominance was deliberately misleading. In this case the authors correctly demonstrate that there are limits on the attraction effect, but do not clearly define the conditions where which it does not work.

Put differently, there are problems with the practical significance of the attraction effect, but these need to be more clearly defined. The current paper needs to move in the following directions.

1. Emphasize the ease of seeing the dominance relationship. How quickly can a person see that the target is clearly worse than the decoy? How unambiguous is that relationship? For example, the price differences are harder to compare on page 13 in the visual contrast. Thus, concentrating on either attribute in a lexical sense may limit the context effect. Further, while the image of the sea is clearly better than the inland one, both have positive and negative aspects so that the dominance is not immediately apparent or even in reflection without counterarguments. The same conflict can occur with taste. Generally, the attraction effect will not occur for attributes that non-monotonic (such as the level of salt or temperature), or with very rich attributes such as cities, or brand names.

2. Consider associations between attributes. Attribute linkages can alter the meaning of attributes. Thus, as noted in the last round, the bad orange could poison the perception of the good one, as could a bad taste of grape Kool-Aid. Attributes with correlated utilities (substitute, complementary attributes) have typically been avoided in tests of attraction effects because they could cause the effect, or make it go away. You now discuss such associations as repulsion effect. However, such associations should be considered as independent of the attraction effect, being able to increase or decrease the effect. For example, the greater impact of high over low quality decoys can be accounted for by the positive association between two high quality items (see again Heath and Chatterjee’s work). Try a pretty good and a very good apple rather than a bad one.

3. Be clearer on the boundaries. It is not just numbers that work, the attraction effect can occur with stars, or clear relative sizes. You make a reasonable point about probabilities, but they are hard to understand and in a quick, non consequential choice, may be ignored. Non-compensatory processing in the face of the decoy can either increase or decrease the attraction effect.

It is not possible to clarify all of these issues, but they should be better raised. Currently, the paper is written as too much of a gotcha on the field accepting results that appear reasonable but are not. Perhaps the gotcha is deserved, but simplifying the problems with the attraction effect can be as reprehensible as simplifying its difficulties.