**RE: “The Rules of Attraction” (JMR-12-0061)**

Dear Professor Erdem,

We thank you and the review team for reading our manuscript, for the opportunity to submit a revision, and for the extension to complete data collection on another round of studies. We recognize that it is conventional here to describe how we modified the manuscript to address the concerns of the review team, which you summarized and amplified. However, as we explain below, we chose not to follow the recommendations, because we disagree with the substance of the objections. Thus, instead of repositioning the paper, we’ve doubled down on what we offered earlier, and will try to strengthen our case here for why this paper is important, since the central concern of the review team is that it is not.

The attraction effect (aka, asymmetric dominance effect or decoy effect) is among the most studied and celebrated phenomenon in the behavioral marketing literature, and is widely perceived as large, robust, and important, as evidenced by the quotes below:

*[We conclude] that the attraction effect is robust, has a wide scope, is quite sizeable and is of practical significance.*

(Doyle, O’Connor, Reynolds, and Bottomley, 1999, p. 225)

*Decoy effects …occur in product classes ranging from restaurants to light bulbs and occur regardless of whether choice sets are manipulated between subjects or within subjects.  [They] are important for both theory and practice*.

(Heath and Chatterjee, 1995, p. 268)

*[T]he attraction effect is a real-world phenomenon, not just an experimental artifact.*

(Mishra, Umesh and Stem, 1993, p.331)

*Asymmetric dominance and advantage (decoy) effects can exert a powerful force on choice because they provide a compelling justification for the purchase of one option over another.*

(Kivetz, Netzer, & Srinivasan, 2004, p. 265)

*[The attraction effect is] a general feature of human choice behavior because [it is] a fundamental part of decision-making processes.*

(Trueblood, Brown, Heathcote, & Busemeyer, 2013, p. 906)

Given these widespread endorsements of the magnitude and scope of the phenomenon, we recognize why a paper that disputes its practical reality would be controversial. What we do not understand is why it is regarded as unimportant or an insufficient contribution.

We are puzzled by the reaction of the Associate Editor and Reviewer 1, because many of the studies we conducted are of exactly the sort called for in prior work on the effect:

“*Future researchers examining context effects on choice behavior… should create more realistic situations.. involving actual choices between physical products…Indeed the resolution of the influence of the meaningfulness of product information on the presence or absence of the attraction effect may not be possible without experiments involving actual products, particularly if sensory information is critical to choice*.”

(Ratneshwar, Shocker, & Stewart, 1987, p. 551)

“..future research should *test whether the attraction and compromise effects and their explanation still hold in more natural consumer environments*.”

(Simonson, 1989, p. 173)

“…*consideration of external validity is essential when studying the construction of preferences. Specifically because the conditions under which preferences are formed interact with the resulting preferences, one cannot extrapolate from studies that misrepresent reality with respect to the construction of preferences in general*.”

(Simonson, 2008 p.260)

Of course, we assume some responsibility for not having done a better job clarifying the contribution, and hope the revised manuscript is improved in this regard. We’ve also conducted several additional studies that strengthen our earlier claims.

In what follows, we offer detailed responses to most of the concerns of the review team.

**AE Report**

**I have some difficulty seeing how each product class was really even a test of the attraction effect in the first place.  For example, with apartments, the decoy only differed from the target in that it offered a view with a cloudy sky.**

As you suggested, we now reproduce the stimuli from these studies in Appendix B; the first set of stimuli reproduces the choice set in question, which involved apartments varying in size (specified numerically in square feet) and view (depicted with photographs). One apartment was small (530 square feet), but had a very attractive view, of what appeared to be the ocean shoreline. The target apartment was larger (910 square feet), but with a much less attractive view. The decoy was dominated by the target on both dimensions – it was slightly smaller (905 square feet) and the view appeared even worse (essentially the same view, through a dingier window). We referenced this as “cloudy sky” in the corresponding table. We understand and apologize for the confusion created by our odd choice of nomenclature.

**Does this make the other option dominate this option? It's really hard to say.**

Relative to the target, the decoy apartment is worse on both of the specified dimensions: it is smaller and has an even worse view. So it seems rather clearly dominated. Moreover, respondents apparently recognize this, because only 2% of them choose it.

**Would someone really take this into account in his or her decision?**

48% of respondents chose the option we intended as target, compared to only 2% who choose the option we intended as the decoy. So it seems safe to conclude that respondents are accounting for this in their decision.

**For fruit, what did the pictures look like?  Were the black spots and dented skin apparent?  Or, were they so significant that they resulted in feelings of disgust and possible contamination of the other orange?**

The pieces of fruit were depicted using color photos, now reproduced in Appendix B. The decoy orange had two spots of discoloration. The decoy apple appeared to be bruised. So, yes, the imperfections of the decoy options were apparent. As intended (and as reported in Table 2), the decoys were essentially never chosen. We can only speculate what feelings the imperfect fruits evoked. To our surprise, their presence had no effect on the choice share between the two “normal” fruits – that is, we found neither attraction nor repulsion.

**If the orange really looked disgusting, and you don't find the attraction effect, this would be like telling a manager to use an inferior TV, but don't bring in a TV that is smashed up with insects crawling out of it.**

We now reproduce the materials in Appendix B. You can judge for yourself whether the orange was disgusting or merely inferior to the more perfect orange. With respect to your question regarding the sale of television sets, if a broken, insect-infested television can be regarded as an analog to a discolored orange, we would tell the manager that adding that option to the product line is unlikely to increase sales of the normal television – a message consistent with the one advanced in our paper.

**For hotel rooms, the decoy was a room with "nice" décor and a polygon bath?  Is a polygon bath really supposed to be inferior somehow?**

Of course, our verbal summaries were impoverished versions of what respondents actually encountered – shortened so as to fit into a table. We now reproduce all of the visual stimuli in appendix B. With respect to the stimuli in this study the presence of a hot tub was one salient indicator of quality. The two expensive rooms (the target and the decoy) each possessed one, the cheaper room did not. Among the two rooms having one, other visual features were intended as cues for quality. We intended for the decoy to appear less fancy than the target option and, hence, to be dominated, since their price was the same ($180 per night). Respondents largely interpreted the stimuli as we intended, since the target was chosen much more frequently (67%) than the decoy (13%).

The repeated objections to our stimuli appear to overlook two key features: (1) the decoy was self-evidently more similar to the target than to the competitor in every case; and (2) the decoy was dramatically less popular than the target in every case, as Table 2 reveals. Thus, a vast majority of respondents are, in fact, appreciating the intended differences: there is no other way to account for the data. Moreover, even when the experimental design was not fully successful (i.e., when the fraction preferring the decoy is non-negligible), this at most creates difficulty in interpreting the change in choice share, and thereby diminishes the evidential weight of that study. It in no way invalidates the general approach or the collective implications of the remaining studies.

**I could go on like this for most of the stimuli used in Table 1.  Thus, before just dismissing an effect, it is important to carefully do this, by making sure that the target really dominates the decoy,**

As noted above, we don’t think that our stimuli are so problematic, and we hope that this is clearer with our inclusion of Appendix B.

**but the decoy isn't so bad that it results in disgust (and to show us exactly what you used as stimuli).**

With respect to the capacity of our decoy options to elicit disgust, we suspect very few met this threshold for most people. None of us are disgusted by a nice hotel room that is slightly less nice than an even nicer one, or by larger apartments with unappealing views, or by weakly flavored fruit drinks, or critically panned movies. Two of us would probably even prefer a flawed orange to a perfect apple. Across the 36 studies we conducted, the decoys elicited disgust in probably just one case – the jellybean study, in which one of the two conditions actually consumed *Bertie Botts* jellybeans flavored to resemble pepper, dirt, grass, and ear wax, respectively.

But, in any case, it is not obvious why the putative presence of disgust is necessarily invalidating or uninteresting. For instance, one could imagine making the image of the decoy orange sequentially less attractive up to, and into, various degrees of disgusting. It is quite possible that the size and valence of contextual effects would respond to such manipulations. That is an empirical question, and a somewhat interesting one, though our studies were not designed to address it. (We should note that in the one study involving unambiguously disgusting stimuli – the jellybean study – the decoys actually had little effect on the choice share of the flavors whose appearance they most closely resembled.)

**I think the basic point that you are really trying to make is that the attraction effect will only hold when it is obvious that one option dominates the other. But, again, is that news?  I'm not so** **sure.**

That is *not* the basic point we are trying to make. As noted above, dominance is quite clear in most cases. We are alleging that the attraction effect appears to be restricted to highly stylized stimuli that involve 2 X 2 numeric matrices.

We base this strong claim on our repeated failure to find such effects outside of these contexts using studies sufficiently powered to detect even moderate effects. In all we tested for the effect 27 times and failed to find it 27 times. We think this yields a conclusion.

**At the same time, I do understand your point that if a manager tries to implement the attraction effect, it is hard to say what will happen in the real world.  Will everyone really see the target as dominating the decoy?  Did you do any field experiments?**

We suppose it depends what you mean by “field.” We did several studies involving actual consumption of actual goods (and are one of the few to have done so). In any case, we see no basis for assuming that effects we fail to find in the lab would be present in the field.

**For study 2, I guess again I'm wondering how this is different from Ratneshwar et al.**

There are two main differences: (1) we are not *supplementing* an ambiguous numeric specification with a somewhat less ambiguous verbal description, but are rather *replacing* a numeric specification (probability) with its perceptual counterpart (the shaded fraction of a probability wheel). (2) We find no significant attraction effect, whereas Ratneshwar et al. continue to find significant effects.

**For study 3A, I looked at Figure 2, and at first glance I could not tell the difference between Picture B and Picture C.  Thus, how can we expect the attraction effect to hold if someone doesn't even notice that the target dominates the decoy (whereas this is obvious with the quality ratings)?**

Though we admit the visual differences are subtle – especially if you do not use a color printer – respondents *can*, apparently, discriminate the two stimuli. Once again, of the 60 respondents in that condition of the study, note that only 1 of them chose the decoy, compared to 21 who chose the target. If respondents truly regarded the picture qualities of the TVs as identical, you’d expect the *opposite* pattern, since the decoy is slightly less expensive ($339 vs. $350).

**And, for study 3B, I'm afraid that you have insignificant results (no attraction effect, no repulsion effect).  Thus, I'm not sure how you can really use Study 3B to build your paper.**

We contend that the attraction effect does not occur if you move away from highly abstract “schematic” alternatives that amount to 2 X 2 numeric matrices. Correspondingly, insignificant results are at the very heart of our case, and, thus, finding another one here adds evidentiary weight to this claim. Going beyond the epistemic value of the null effect reported for the “choose first” condition, the contrast between this null effect and the (now significant) effect in the “rate first” condition suggests that contextual effects can be influenced by supplementing perceptual representations with (self-generated) numeric representations. However, this study was only intended to explore the contours of the boundary conditions for the effect, the results of this “rate first” condition do not bear strongly on our claims.

\*We note that in the revised manuscript, the study in question is now 3C rather than 3B.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Reviewer: 1**

**The main proposition of the current manuscript is not so interesting, not well supported, and not original.**

We will address these three criticisms in turn.

**Not so interesting**

There is no arguing with tastes, we suppose, but we will say three things in our defense:

**(1)** Ratneshwar, Shocker, and Stewart (1987) and others after them (see, e.g., Simonson 1989), have urged the field to test whether attraction effects occur in more natural consumer environments – precisely what we do here.

**(2)** We opened our rejoinder with a sample of quotes to illustrate how the attraction effect is discussed. To us, these quotes represent widely distributed false beliefs among both academics and practitioners. If a false belief becomes sufficiently prevalent, it eventually becomes important to show that it is false, even if the truth is less interesting than the belief.

**(3)** Your reaction may not be typical. Reviewer 2 characterizes the work as “refreshing.”

**Not well supported**

We hope you find that the revised manuscript provides greater support for our position.

**Not original**

We think our work is good ordinary science and will not try to exaggerate its originality. We are certainly not the only researchers to have used more realistic stimuli when investigating contextual effects (though across three decades of writing on the topic there are remarkably *few* such studies; see appendix A). However, two things are worth noting.

**(1)** We agree that Ratneshwar, Shocker, and Stewart (1987) were one of the first to examine an issue we also examine here – boundary conditions for the attraction effect. However, our manipulations, results, and claims differ from theirs in the ways we note above, in our response to the Associate Editor regarding study 2. Furthermore, if experimental stimuli were arrayed on the continuum ranging from extremely abstract/highly stylized to fully realistic, theirs would lie almost as far to the left as the typical study on this topic, whereas most of ours lie pretty far to the right.

**(2)** By drawing upon much more data and a much wider variety of designs, we’re able to make a much bolder claim – that the attraction effect is nonexistent in many or most realistic contexts. In addition to the study designs we conceived, we reexamine the small literature which reports significant effects in less stylized settings. In each case, we failed to replicate the reported results.

**Although it is true that complex stimuli where the relations among options are hard to detect are less conducive to the attraction effect, the basic effect is still there.**

**(1)** As noted earlier in our responses to the Associate Editor, we disagree that the relations are difficult to detect. We have little doubt that respondents can and do make such judgments; indeed, we suspect that they could often recognize the dominance structure *more* quickly in several of our studies than with the more typical *stylized* stimuli (e.g., that after tasting three samples of Kool-aid, participants likely recognize that dilute grape tastes a lot like regular grape, except less flavorful).

**(2)** Regarding the assertion that the “basic effect” is “still there,” we respectfully disagree. Of course, we can never prove that there are no black swans outside the lab, but across 27 studies we fail to find a single one – not even when we look in the few places in which previous researchers have claimed to find them. We think that is enough data to draw a practical conclusion, to initiate a dialogue about the practical reality of the effect, and to promote further discussion about its boundary conditions.

Of course, even those who are prepared to concede that black swans are absent in the wild might reasonably remain interested in how swans can be made black in the laboratory. Our studies speak to these people as well. Importantly, since the effect appears to require that at least two attributes be numerically denominated, our results support *tradeoff contrast* (Simonson & Tversky, 1992) as the mechanism causing the blackness of swans in the lab. Correspondingly, our results weigh strongly against other interpretations of the effect that others have championed, because none of these other mechanisms would explain the striking disparity between perceptual and numeric representations of product attributes.

**Reviewer: 2**

**You define an important boundary condition for the attraction effect. Any researcher who has tested such context effects has found such reversals, but too often not mentioned them.  The broad empirics of your paper are therefore refreshing and believable, and the demonstration of a repulsion effect could be quite important.  Below are a number of suggestions for your paper.**

Thanks for the kind words about our paper. We hope you find the revision even more compelling (even if we didn’t follow all of the promising leads you proposed). We address some of your comments below.

**1.  Be clearer on your sampling scheme for Study 1. One of the problems with the choice-context effects literature is that choices across product classes are easily run.  It is thus easy to ignore those that do not fit the expected effect, leading to results that seem reliable but may be spurious. Accordingly in your table 1 it is important to be clear on the set of product categories tested. Did you have, say, 100 different tests of attraction, most of which used numbers, but the rest did not?   Alternatively, Table 2 could reflect a biased sample of the non-numerical tests performed.  I assume that was not what you did, but clarity is needed**.

We apologize for the confusion on this point and hope it is clarified in the revised manuscript. Table 1 is, in fact, exhaustive. It lists every study we’ve ever run that is not listed elsewhere in the paper. In other words, there was no selection. Essentially, the existing literature was the control group for studies 1a-1s. The existing literature shows many examples of sizable attraction effects using stylized stimuli (2 X 2 numeric matrices). Heeding earlier calls for more research on contextual effects outside of this paradigm, we constructed choice sets that would retain the essential elements of asymmetric dominance (i.e., a core set compared to an expanded set that included a decoy which was similar, but inferior to one of those options). Our intent was to test under which situations (if any) the attraction effect would replicate. We found no evidence of the effect in any of the studies. As noted above with the black swan metaphor, this does not of course mean that no effect *could* be found, but it is certainly very informative regarding the boundary conditions of the effect. Our other studies (2a, 2b, 3a, 3b) and our newly added Appendix C (which outlines our failures to replicate several published results) substantially strengthen the collective implication of the (non) results from Studies 1a-1s

**2.      Focus on the cases of significant repulsion effects.  Null effects are OK, but less meaningful theoretically.  Additionally, anything that makes dominance hard to immediately recognize should reduce the attraction effect, but that is far less interesting than the repulsion case. Thus emphasize repulsion more.**

Using stimuli with greater realism, which involve perceptual representations of one or more attributes, we repeatedly find that the decoy either fails to affect (or even diminishes) the choice share of the target option. These results certainly have theoretical meaning, however they are labeled. When viewed under the assumptions of early choice models (e.g. those by Luce), they might rightly be called *null effects*, contrasted with the attraction *effects*. However, when viewed against the backdrop of studies finding attraction effects with stylized stimuli, our results might be considered *effects*, since perceptual representations could be seen as *affecting* choice processes: whether by facilitating some process that is absent with stylized stimuli or by impeding some process that is otherwise present (such as tradeoff contrast).

The common inclination to dismiss so-called “null” effects contributes to the problem you cite in point 1, in which the published results create a distorted representation of reality. Existence proofs of a psychologically interesting phenomenon were a sufficient justification for many of the first publications on the attraction effect, but the message that academics and practitioners now draw from their reading of the literature has been grossly deformed by a combination of file drawer effects, false positive results, and an inadequate attention to boundary conditions. Our paper attempts to correct this. We regard this as a substantial and sufficient contribution.

**3.      Build up a theory of repulsion.  It seems most likely when the decoy's attributes reinforce an association (e.g., dilution with either grape or cherry flavor, duck farts with spring water) and that association generates an aversion to similar stimuli.  A way to start would be to rank the size of the aversion effect in Table 2 and build up hypotheses about what is driving the results found.**

We set out to delineate the boundary conditions of the attraction effect, and sort of stumbled across occasional instances of this opposing phenomenon, which we term the *repulsion effect*. These are intriguing, and a paper exploring the repulsion effect may be very interesting. However, though we understand that this is the paper that you (and the AE and the editor) want us to write, it is not the paper we want to write.

Though we admit that a sub-goal of our paper *is* to stimulate further research on the repulsion effect, for the purposes of this paper, we are mostly content to simply treat significant repulsion effects as especially strong evidence against attraction effects. We are not yet prepared to say much more and further attempts to do so would likely distract from our central message regarding the practical reality of the attraction effect.

Regarding repulsion effects, the account you offer above is essentially identical to our own thoughts on the issue. However, we will note that a casual version of such an analysis does not yield the patterns one might expect. Namely, we fail to find repulsion effects in contexts in which one might expect the aversive decoy to “taint” the most similar members of the choice set (i.e., in studies involving the consumption of disgusting jelly beans and pictures of damaged fruit).

**4.      You argue appropriately that both relative tradeoff rates and range and frequency effects will be less apparent when numbers are not included.  However, as Wedell and Pettibone demonstrate, the major driver of attraction is quickly seeing that the target dominates the decoy.  You need to test for the ease by which that dominance is perceived.**

It is not so much that perceptual stimuli make tradeoff rates less *apparent*, but rather that tradeoff rates will be impossible to compute unless both dimensions are numeric. Regarding the detection of dominance, it seems clear that respondents *are* detecting dominance for the reasons already outlined in our response to the AE. However, a central message of the paper would actually remain intact if failure to detect dominance *was* the reason we fail to find evidence for the attraction effect. Namely, if dominance were always obscured in realistic settings, then the attraction effect would never occur in realistic settings.

**5.      Be clearer on the boundaries of the attraction effect. Attraction clearly occurs with quasi-numeric quality ratings such as star ratings.  However, does it occur with clearly ordered phrases such as 'top rated' vs. 'consistently good quality?'  You give an example of a fruit with blemishes as the decoy.  The problem there is that one might infer the target had similar unseen blemishes.  Would attraction also work fruits of different sizes?  Put differently, visual images of volume or length may act like numbers. Can you test whether they do?**

Though we think we’ve drawn a clear distinction between situations in which attraction effects are moderately common (choice options represented as 2X2 numeric matrices) from situations in which they are not *at all* common (pretty much anywhere else), further exploration along the boundary might well be warranted, and you draw some interesting test cases of precisely where that boundary falls. Note that in our response to Reviewer 1, we noted a distinction 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 (as in several of our studies).

Similarly, using your example, you can imagine cases in which the numeric vs. perceptual distinction may be blurred. Consider a choice between two packs of postage stamps: one in which the number of postage stamps is specified and one in which the stamps are “merely” shown, but readily countable. Would these two representations yield identical choices and identical contextual effects? We are unsure. In studies 2a and 2b, we *do* find considerable differences when we compare numeric and visual representations of probability, though we do not know whether this would extend to some of the test cases you suggest. We’d conjecture that the visual representation would not yield an attraction effect, unless the stamps were in fact, counted and overtly recorded, akin to our study 3c. We’d hope to avoid having to submit an exhaustive analysis of the precise contours of the boundary we emphasize, though we acknowledge that further work along these lines is very much in the spirit of the paper we want to submit.

**6.      Study 2a is really a fine study, with many respondents and 2b provides a really clear case where dominance is readily apparent.**

Thank you. One of the strengths of our manuscript is the large number of respondents we use to obtain a more reliable estimate of the size of the effect. We typically fail to find attraction effects despite these comparatively high powered studies.

We agree that dominance is readily apparent in 2b – and arguably even more salient in 2a, in which the decoy is dominated on *both* dimensions.

**7.      Study 3b is intriguing, implying that asking for numerical rating of attributes provided by the chooser will resurrect the attraction effect.  How important is a numerical response? Would the same reversal occur if respondents merely were asked to express their attribute valuations on a non-numerical line?**

That is a very interesting question. Indeed, we expanded this study (now study 3c) so as to test a variant of your idea. In that study, ratings on a visual analog scale acted similarly to numeric ratings, and, thus, we report only the pooled results (see endnote #8). However, in that study there were no large contextual effects in *any* condition, so a conclusive answer to your interesting question remains out of reach for now.