Zero attraction: A precisely zero attraction effect with naturalistic stimuli that meets Huber, Payne, and Puto's (2014) criteria

Anna Trendl

Department of Psychology, University of Warwick

Neil Stewart and Timothy L. Mullett

Warwick Business School, University of Warwick

November 19, 2018

Abstract

In the attraction effect, adding a third option to a choice set of two options can reverse the preference for original options and even increase the choice share for one of the original options. For this reason, the attraction effect is theoretically important, as it is a violation of the properties of regularity and independence from irrelevant alternatives, which are core properties of any model in which the utility of each option is stable across choice sets. The attraction effect has driven the development of over X Hopefully X is 10 models of multiattribute choice in the past 20 years. Recently, ? (?) claimed that the attraction effect is limited to options presented as collections of numerical attributes, and does not hold for choices between naturalistic options like snacks or movies—a claim which would severely undermine the importance of the effect for the development of models of choice. ? (?) criticised ?'s experiments, laying down a set of strict criteria that a naturalistic attraction effect experiment should meet. We present the first experiment that meets these criteria. We find a precisely zero attraction effect.

Neil has comments like this

Anna has comments like this

Tim has comments like this

1   To Do

1. Add in text explaining why the attraction effect is important: The attraction effect is important because it rules out the simple scalable family of models, where the value of each alternative is independent of the choice set in which it appears. For the introduction of the decoy to reverse preferences between the target and the comparison is a violation of independence of irrelevant alternatives and often of regularity, where the introduction of a new alternative to a choice set cannot increase the choice share on any existing alternative. Because of its theoretical importance, it has played a significant role in the development of multialternative models of choice, with a significant number quite different theoretical accounts being developed in the past 20 years cite them all, from MDbS. For this reason, it would be of considerable importance if the attraction effect is confined only to numerical stimuli presented in an attribute by alternative format. Arguably, while choices in the real world involve stimuli are often presented in a rich naturalistic form (Bhatia and Stewart, 2018), most experiments which have been used to develop theory use the numerical attribute x alternative format. There are substantial differences in decision making with these naturalistic stimuli. For example, Bhatia and Stewart (2018) show that the heuristic chice strategies that are so well supported by the numerical attribute x alternative experiments are substantially outperformed by a simple weighted additive regression when naturalistic stimuli are used.

2. Highlight that although the stimulus set construction is weird, it is all validated by the ratings then choice data from the experimental session after the preregistration. Something upfront like "While the details of our stimulus construction are necessarily somewhat arbitrary, this does not undermine the quality of the experiment. We were able to check, using only the responses from the experimental session after stimulus

construction and preregistration that our triplets meet every one of the the Huber et al. requirements."

3. Add the list of stimuli, the preregistration, the data, and the R script to an OSF page so we are sharing everything up front. Don't make public yet.

2 Introduction

Imagine that you are facing a choice between two hotels for a weekend city break: Hotel A has an excellent central location but is fairly expensive, whereas Hotel B is located in the outskirts of the city, but has a very favourable price in comparison to Hotel A. You are torn between the two hotels, when you spot a third option: Hotel C is situated right next to Hotel A in the city centre, but is slightly more expensive. Which one would choose?

Decades on decision-making research suggests that the presence of Hotel C (decoy) in the choice set makes you more likely to choose Hotel A (target) over Hotel B (competitor). This phenomenon is called the attraction effect (also known as asymmetric dominance effect), and it is one of the most widely researched decision biases in consumer decision making. In essence, it states that when the decision maker is indifferent between the target and the competitor (Hotel A and Hotel B in the example), the addition of an inferior decoy option that resembles the target (Hotel C is similar to Hotel A, but is slightly worse) increases the likelihood that the target will be chosen.

The attraction effect is important because it rules out the simple scalable family of models, where the value of each alternative is independent of the choice set in which it appears. For the introduction of the decoy to reverse preferences between the target and the comparison is a violation of independence of irrelevant alternatives and often of regularity, where the introduction of a new alternative to a choice set cannot increase the choice share on any existing alternative.

The real-world relevance of the attraction effect in consumer decisions has recently become a contentious issue in the decision-making literature. In their study, ? (?) present

a thorough investigation of the boundary conditions of the attraction effect based on 38 experiments with various stimuli types. These stimuli types include choice options with numerically represented attributes as well as complex, real-world stimuli (e.g., fruits, bottled water, apartments, etc.), and in some of these experiments participants could even sample the choice options (e.g. squash, mints, popcorn). The overall conclusion of this study is that while the presence of the decoy seems to affect decisions when the option attributes are represented numerically, it is absent in experiments with more complex, naturalistic stimuli. In light of these results, ? posited that the psychological processes underlying decisions that involve options with numeric attributes are fundamentally different from those employed in decisions where the stimuli has a naturalistic representation. This conclusion was also supported by ? (?), who reported difficulties replicating the attraction effect when the stimuli were pictorial, as opposed to when attributes were presented numerically.

These two studies sparked considerable interest amongst decision making researchers, and led to the re-examination of the boundary conditions of the attraction effect. ? (?) discussed five critical conditions that can inhibit the attraction effect, and argued that many of these are present in the experiments reported by ? and ?. These are the following: (1) strong prior preferences over the target and competitor, (2) inability to identify the inferiority of the decoy, (3) heterogeneity in prior preferences over the target and competitor, (4) an undesirable decoy and (5) a too desirable decoy. ? (?) further stressed the importance of the detection of the dominance relationship in observing the attraction effect, and also pointed out several other smaller, specific methodological shortcomings of the studies by ? and ?.

Owing to its theoretical importance, the attraction effect has played a significant role in the development of multialternative models of choice, with a significant number quite different theoretical accounts being developed in the past 20 years cite them all, from MDbS. For this reason, it would be of considerable importance if the attraction effect is confined only to numerical stimuli presented in an attribute by alternative format. Arguably, while choices in the real world involve stimuli are often presented in a rich

naturalistic form (Bhatia and Stewart, 2018), most experiments which have been used to develop theory use the numerical attribute x alternative format. There are substantial differences in decision making with these naturalistic stimuli. For example, Bhatia and Stewart (2018) show that the heuristic chice strategies that are so well supported by the numerical attribute x alternative experiments are substantially outperformed by a simple weighted additive regression when naturalistic stimuli are used.

In this article, we describe our attempt to test the attraction effect with complex, naturalistic choice options, using an experimental methodology that addresses the issues raised in connection with the work of ? and ?. Using a choice task where we ask people to choose between movies, we find no evidence for the attraction effect, that is, the presence of the decoy does not affect preferences between the target and the competitor. We discuss the implications of our findings for understanding the boundary conditions of the attraction effect.

3 Testing the attraction effect with real-world stimuli

The goal of this experiment was to test the attraction effect with naturalistic stimuli. We chose to use the most popular movies on IMDb as stimuli for several reasons. First, since movies are an integral part of Western culture, we can reasonably expect that most our participants will be sufficiently interested in the choice options. Second, choosing between movies is a realistic, everyday task, and this way we can mimic real-world decision making. In addition, the movie space is rich enough so that we can create a wide range of unique choice triplets to test the attraction effect.

When the stimuli has numerical attributes, it is straightforward to construct choice triplets with a target, competitor and decoy. However, with naturalistic stimuli, this task is significantly more complicated. ? have also used movie stimuli in two of their experiments: they chose pairs of movies that are part of the same sequel or are starring the same actor (but have distinctly different genres) to create target-decoy pairs (they

Figure 1: A quadruplet made up of two attraction effect triplets. Triplet 1 consists of Movie A (target), Movie C (competitor) and Movie B (decoy to Movie B), while Triplet 2 consists of Movie C (target), Movie A (competitor) and Movie D (decoy to Movie C).
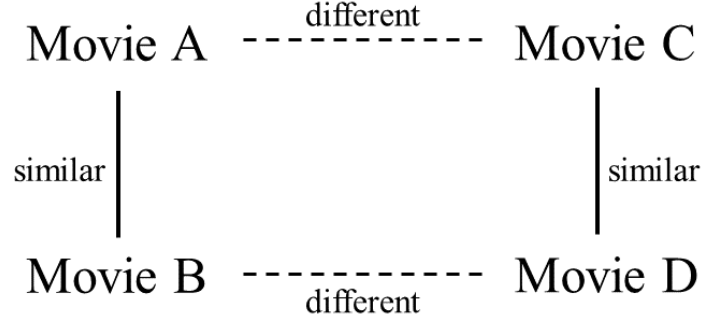
used eight distinct movies overall). In the movie sequel version, they assumed that the second part of the sequel is the decoy, and the first part is the target, whereas in the movie actor experiment it is unclear how they decided which movie of the same actor should be the target and the decoy. We believe that this experimental approach has a significant flaw.

Specifically, there is no guarantee that the participants perceived the triplets as they were intended by the authors: whether *Grease 2* will be perceived as inferior to *Grease* (or the other way around) is completely up to the decision maker's preferences, which they did not control for. Individual heterogeneity in preferences can severely alter the perception of real-world choice triplets, thus, it is imperative to take it into account in the choice triplet selection.

Our novel experimental design takes individual preferences into account and ensures that decision makers are indifferent between the target and competitor and that the inferiority of the decoy is clearly identified. In addition, to increase the statistical power of our test, we used a within-subjects design as well as both A, B, A' and B, A, B' triplet pairs (where X' is the dominated option), as suggested by ? (?). We created "quadruplets" (each made up of two attraction effect choice triplets), using two distinctly different target-decoy pairs. Figure 1 shows the structure of one such quadruplet, where Movies A-B and Movies C-D are the two target-decoy pairs.

3.1 Method

3.1.1 Candidate choice set selection

Our movie triplets are listed in Table X in the Appendix. The details of the construction of these triplets are somewhat arbitrary—a different recipe could have been use. However, the key issue is that our triplets pass ?'s (?) criteria, as we detail below.

We first retrieved the most popular 40 movies from each of 10 distinct genre categories (romance, drama, sci-fi, thriller, comedy, horror, animation, fantasy, crime, action) from IMDb, so that we had 400 movies overall. We chose a wide range of genres to obtain a sufficiently rich stimuli space. We omitted an sequels.

Owing to the multidimensional nature of the stimuli, one of the main difficulties in creating attraction effect choice triplets from real-world objects is establishing a criterion for matching up similar objects. We used genre and sub-genre information from allmovie.com to create target-decoy pairs with overlapping genres that are likely to be perceived as similar and to create target-competitor pairs with non-overlapping genres that are likely to be perceived as different.

We conjectured that it will be harder to find movie pairs that will be perceived as similar because any given movies is similar to a only a few movies and dissimilar to all of the others. For this reason, we started the quadruplet creation with selecting potential target-decoy pairs.

The genre information on allmovie.com is very rich: compared to the 18 genre categories on IMDb, there are 156 genre and sub-genre categories, capturing many important aspects of the movies. Using this rich genre information, we created a movie by movie (400×400) matrix, where each cell was the number of overlapping genre categories between the two movies. We select as target-decoy pairs all of the pairs where, for a given target movie, the number of overlapping genres with a candidate decoy was equal to the maximum overlap seen for that target across all candidate decoy movies, creating 2,271 target-decoy candidates. We also added 806 movie pairs obtained from the mutually closest 10% of movies based on a latent semantic analysis[1] that were not already in our list of

---

[1]The latent semantic analysis assesses the similarity of two items based on the text associated with

target-decoy candidates. The rationale behind using semantic proximity as an additional criterion was to capture movie pairs that are very close to each other in terms of the story themes, but are not the closest on the genre dimension. Overall, we had 3,011 unique target-decoy candidate pairs at this point.

We then reduced the size of this list by selecting the most similar movie pairs. This was done manually by two researchers, who gave a similarity rating between 1–7 for each movie pair (1-least similar, 7-most similar). We only kept the movie pairs that had a similarity rating above 1, weeding out the movie pairs that were obviously not similar, resulting in 1,242 target-decoy candidates. We then divided the 1,242 pairs into six groups of 207 pairs and ran a pilot study where we asked 60 participants to rate the similarity of a randomly chosen group of movie pairs, obtaining 10 independent similarity ratings for each of the 1,242 target-decoy candidates. Participants rated the similarity of each movie pair on the 1–7 scale which also included a "don't know" option. Figure 2 shows the distribution of the average similarity ratings for each movie pair.

Figure 2: The distribution of the average similarity rating for each target-decoy candidate pair (N = 1242). The dotted line is the cut-off value for accepting a movie pair as similar.
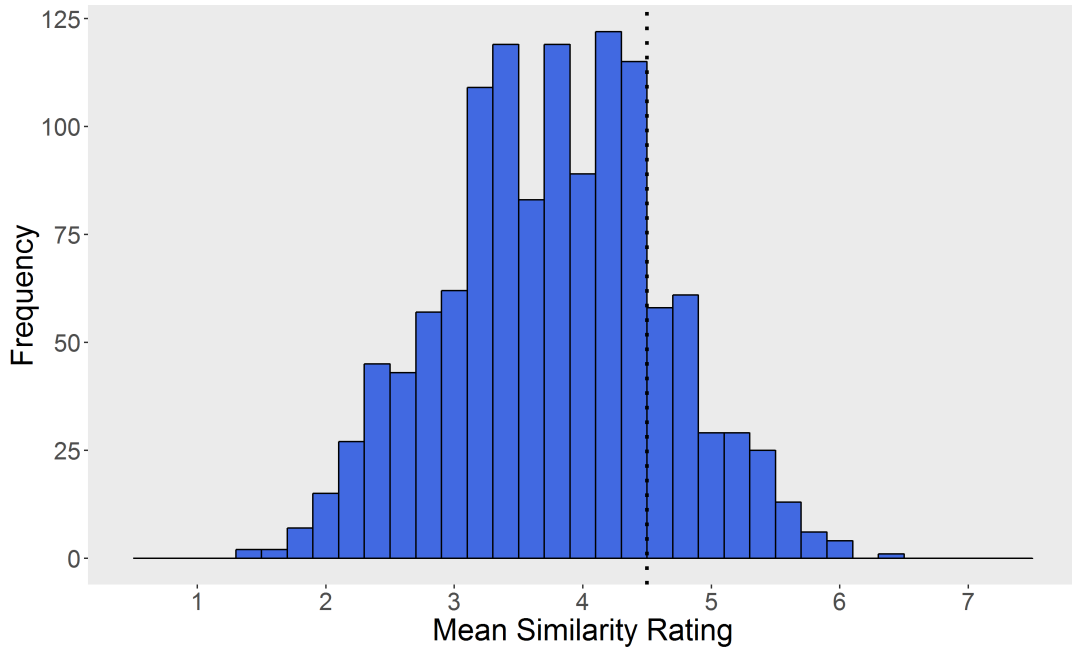
We decided to only use movie pairs with ratings that are equal to or higher than 4.5,

---

them. For this analysis, we used the summary text about the movies as well as plot keywords, actor and director names, all retrieved from IMDb.

which corresponds to the upper 20% of the similarity rating distribution (253 movie pairs).
We intended this procedure will ensure that our target-decoy pairs will be perceived as
similar by most people.

The next step in creating the quadruplets was to create target-competitor pairs. Since
each quadruplet consists of two target-decoy pairs (Movies A and B and Movies C and D in
Figure 1), we decided to pair up the 253 target-decoy pairs to create the quadruplets. To
do this, we first created a target-decoy pair by target-decoy pair matrix (253x253), where
each cell was the number of overlapping genre categories between the two movie pairs.
For example, considering the comparison between target-decoy candidate 1 (consisting of
Movie A and Movie B) and target-decoy candidate 2 (consisting of Movie C and Movie
D), we summed the number of genre overlaps between movies A-C, A-D, B-C and B-
D. We then selected the unique target-decoy pairs that had no genre overlap with each
other. This resulted in 20,022 quadruplets, each of which is a combination of four movies,
created from 231 unique movies. The quadruplets used for each participant are based
upon their own ratings of the 231 movies, as we describe below.

3.1.2 Experimental Procedure

The experiment consisted of three stages: rating stage, choice stage, similarity rating
stage. In the rating stage, we asked for participants' subjective evaluations over the 231
movies ("How do you personally rate this movie?") on a scale from 1 (worst) to 7 (best).
We also asked whether the participant had seen the movie before. The 231 movies were
presented in a random order for each participant. The rating stage took about 15–20
minutes.

Before the choice stage, we created a bespoke set of movie triplets for each participant
using their ratings from the rating stage. First, based on the ratings from the first stage,
we identified the subset of quadruplets where: (a) the target and competitor were both
rated 4, both rated 5, both rated 6, or both rated 7, and (b) the two decoy movies were
rated at least 3 points lower than the two decoy candidates. Note that we did not require
the two decoys in the quadruplet to have the same rating as it would have severely limited

the number of quadruplets we could use (e.g., we allowed for quadruplets with ratings 7,7 for the two targets and 4,1 for the two decoys), but we controlled for this difference in our analysis.

We then selected the subset of quadruplets where all of the movies had been seen or none of the movies had been seen, to make sure that choice behaviour will not be governed by differences in familiarity with the movies. The result was a bespoke subset of quadruplets for each participant where the target/competitor movies had the same rating and the decoy movies were rated worse. However, we did not want the same movie to appear twice as a target/competitor for one participant, and for this reason we used a sequential elimination technique: we first chose the quadruplet with the highest combined target-decoy similarity rating, then eliminated all quadruplets with the same target/competitor movies. We repeated these steps until we had a set a of quadruplets with unique target/competitor movies.

We only invited those participants back for whom we could create at least three unique quadruplets (corresponding to at least six attraction effect choice triplets). In the choice task, people were presented with the selected movie triplets in a random order and were asked to choose the one they liked the most (see Figure 3).

Figure 3: Choice stage.

For you personally, which movie is best?
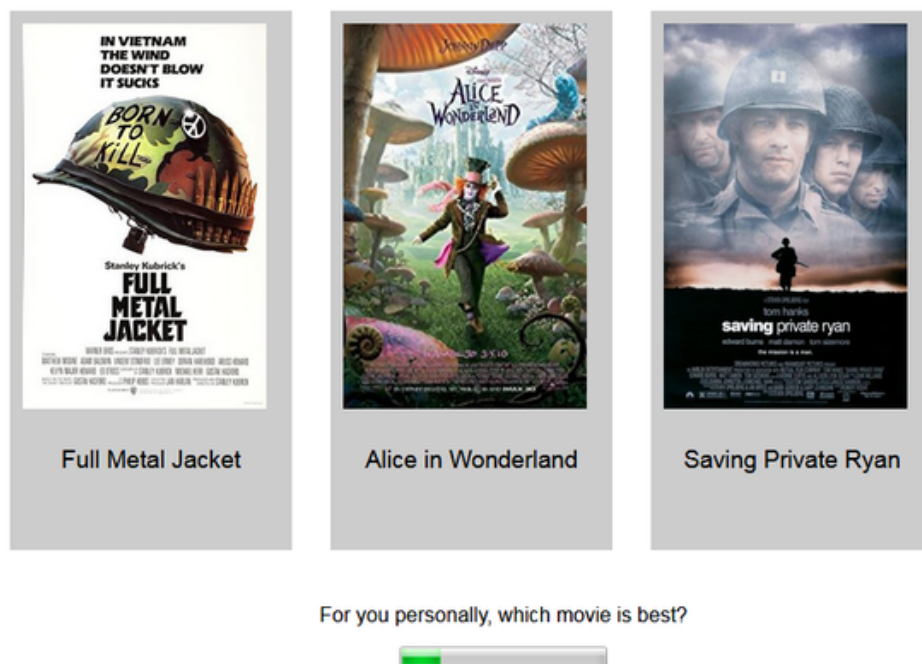
In the final, similarity stage, we asked participants to rate the similarity of all target-decoy and target-competitor pairs on a scale from 1 (least similar) to 7 (most similar), where a "don't know" option was also included. Information collected in this similarity rating stage was important to check that out choice participants each perceived the similarity between the movie pairs as we intended.

We collected data in batches of 50 until we had choice data for at least a 100 participants (after all the exclusion criteria had been applied, see section 3.1.3). At most a few days passed between the rating and choice stage. We recruited 297 participants from Prolific Academic who were paid £8 per hour. Out of the 297 participants that completed the rating stage, we could create quadruplets for 179 of these participants. Out of the 179 participants who were invited back, 152 took part in the choice stage of the experiment.

3.1.3 Exclusion criteria

To conduct a rigorous test of the attraction effect, it is crucial that people take the task seriously and reveal their true preferences. Given that individually rating 231 movies can seem somewhat mundane, we specified a set of exclusion criteria to filter those people out who did not take the rating task sufficiently seriously. These were the following. We excluded people who fell into the fastest 5% of the reaction time distribution, the lowest 5% of the entropy distribution and the upper and lower 5% of the autocorrelation distribution. Entropy refers to the diversity of the ratings, while autocorrelation takes into account the temporal pattern and measures the extent to which a response depends on previous responses. Thus, this procedure filtered out response patterns where people (a) spent an unusually short time completing the task or (b) did not use the whole of the ratings scale or (c) often gave the same ratings for consecutive movies or (d) were giving ratings randomly.

These exclusion criteria were validated in a pilot study, where we collected ratings for a set of books and included repeat trials. The participants filtered out by these three criteria are the ones who give the least consistent ratings to repeated stimuli (correlation

$r < 0.8$).

We applied these exclusion criteria to our sample of 152 people, excluding 17 people, leaving a sample of 135 participants.

The study design, exclusion criteria and all the analyses were planned and registered before we collected any choice data. The pre-registration can be accessed at Insert link to OSF page.

3.1.4 Results

The average number of choice trials per participant was 16 (the lowest being 6 and the highest 54), and 84% of participants were presented with at least 8 choice trials.

The decoy was only chosen in 4.3% of the trials. In addition, 72% of the participants have never chosen it, and only 2% of people have chosen it in more than 25% of the trials. This indicates that people were able to identify the dominated decoy in the choice stage.

To test for the presence of the attraction effect, we first excluded all trials where the decoy was chosen and then conducted a one-sample t-test to test the hypothesis that the mean of the proportion of trials where the target was chosen is above 0.5 (indicating an increased likelihood of choosing the target item). We found no evidence for the attraction effect (M = 0.5, SD = 0.07), t(134) = -0.44, p = .669. Figure 4 shows the distribution of the proportions of trials where the target was chosen, which shows that participants were almost perfectly indifferent between the target and the competitor, M = 0.5, 95% CI [0.49, 0.51].

As specified in our pre-registered analysis plan, we also ran a mixed effects logistic regression with subject-specific intercepts to investigate how the target-decoy and target-competitor similarity ratings and familiarity of with the movies affect the likelihood of choosing the target. Table 1 shows the results from this regression, where contrary to our expectations, none of the explanatory variables seem to modulate the strength of the attraction effect.

While we crated the choice triplets very carefully to ensure that they actually represent an attraction effect choice situation where the target-decoy pairs are perceived as

Figure 4: Proportion of trials where the decoy was chosen. Each dot is a participant and the red dot and error bars show the bootstrapped mean and 95% CIs.

Table 1: Odds-ratios and 95% CIs from a mixed-effects logistic model with subject-specific intercepts. (T - Target, C - Competitor, D - Decoy)

	Dependent variable:
	Target chosen
Seen all	1.122 (0.731, 1.729)
TC similarity rating	0.967 (0.872, 1.073)
TD similarity rating	0.926 (0.835, 1.027)
TD rating difference	0.935 (0.837, 1.044)
Intercept	1.062 (0.598, 1.882)
Observations	1,541
Log Likelihood	−1,064.822
Akaike Inf. Crit.	2,141.644
Bayesian Inf. Crit.	2,173.686
Note:	*p<0.1; **p<0.05; ***p<0.01

similar and the target-competitor pairs perceived as different, individual heterogeneity and malleable preferences can still affect the results. Therefore, we find it instructive to plot the target-decoy and target-competitor similarity rating distributions. Figure 5 shows these distributions, and we can see that the overwhelming majority of target-competitor pairs were perceived as not similar, while the majority of target-decoy pairs were perceived as similar, as we hoped.

Figure 5: Distribution target-competitor and target-decoy similarity ratings.
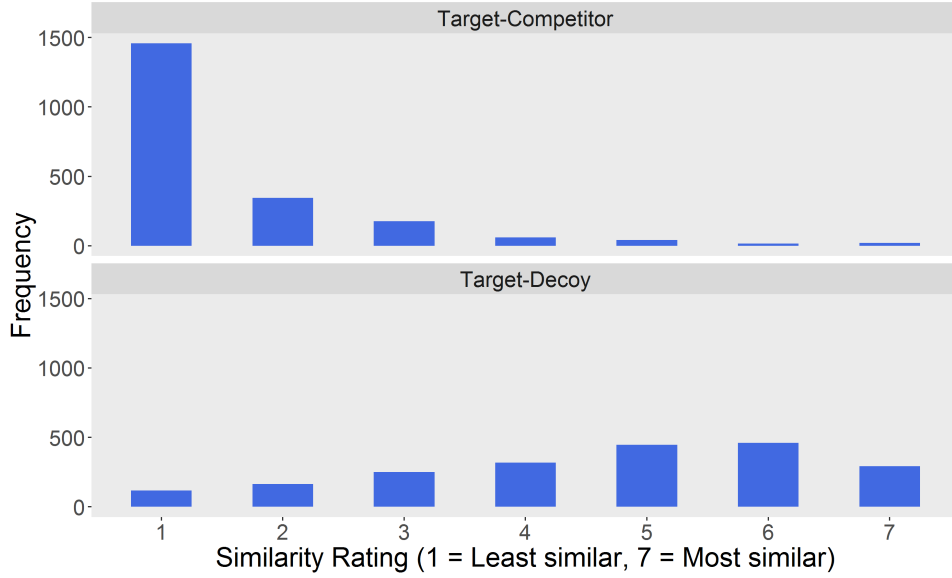
4 Discussion

We tested for the attraction effect in consumer decision making using naturalistic stimuli. We found that the presence of the decoy in the choice did not alter preferences over the target and the competitor, as participants remained indifferent between the two dominating options. We believe that our study is the first rigorous investigation of this research question. We designed our study carefully to address all the criticisms raised in connection with the study by ?, where they used similar stimuli.

First, our experimental design ensured participants' indifference between the target and the decoy, maximising the probability that choices will be constructed on the spot (rather than through relying on strong prior preferences) and an attraction effect will occur. While one could argue that mnemonic processes arising from familiarity with the stimuli can alter preferences in the choice stage, we still did not detect an attraction effect when participants were not familiar with the movies.

Second, we have strong evidence that the dominance relationship was perceived in our experiment. The target-decoy similarity ratings confirmed that our careful target-decoy selection process indeed managed to produce movie pairs that were perceived as similar. In addition, we ensured that the decoy was always rated at least 3 units lower than the target (and the competitor). Consequently, the decoy was only chosen in 4.3% of the

trials, which clearly shows that participants were able to spot and avoid the dominated alternative.

Third, by creating bespoke choice triplets based on ratings, we avoided individual heterogeneity in preferences to act as a potential confound. In addition, we ensured that the decoy is not too desirable in comparison to the target. We also used a strict exclusion criteria to filter out participants who did not take the task sufficiently seriously, and with an average of 16 choice trials per participant we avoided participant fatigue.

Finally, in our analysis we controlled for familiarity with the choice options, perceived similarity of the target-decoy and target-competitor pair, and relative preference between the target and the decoy, but we found that none of these modulated the attraction effect.

To conclude, our results corroborate the findings of ? and ? in that we did not find evidence for the attraction effect with complex stimuli.

Contrary to ?, we do not think that this null result undermines the importance of the attraction effect in decision making research. On the contrary, exploring the relationship between attribute characteristics and the strength of the effect can help us understand how the interaction of the attribute dimensions affect the comparison process which bears important implications for formal models of choice.

write about how it is not only present in decisions that involve options with numerical attributes

and then try to conclude that when there are only two distinct/salient dimensions (need not be numerical) that **can be processed independently** it is more likely to appear

however, when the dimensions interact such that independent processing is not possible we don't see the effect