Context effects in similarity judgments

James M. Yearsley[1]
Emmanuel M. Pothos[1]
Albert Barque-Duran[2]
Jennifer S. Trueblood[3]
James A. Hampton[1]

Affiliations and email
1. Department of Psychology, City, University of London, London EC1V 0HB, UK.
James.Yearsley@city.ac.uk; Emmanuel.Pothos.1@city.ac.uk; J.A.Hampton@city.ac.uk
2. Department of Computer Science, Universitat de Lleida.
albert.barque@udl.cat
3. Department of Psychology, Vanderbilt University.
jennifer.s.trueblood@vanderbilt.edu

OSF link for the project: https://osf.io/z5ngk/
Running head: context in similarity; word count (abstract, main text, excluding figure captions): 3182

Abstract

Tversky's (1977) famous demonstration of a diagnosticity effect indicates that the similarity between the same two stimuli depends on the presence of contextual stimuli. In a forced choice task, the similarity between a target and a choice, appears to depend on the other choices. Specifically, introducing a distractor grouped with one of the options would reduce preference for the grouped option. However, the diagnosticity effect has been difficult to replicate, casting doubt on its robustness and our understanding of contextual effects in similarity generally. We propose that the apparent brittleness of the diagnosticity effect is because it is in competition with an opposite *attraction* effect. Even though in both the similarity and decision-making literatures there are indications for such a competition, we provide the first direct experimental demonstration of how an attraction effect can give way to a diagnosticity one, as a distractor option is manipulated.

Keywords: similarity, diagnosticity effect, attraction effect

1. Introduction

Similarity is a fundamental process, relevant to the way we form categories (Medin & Schaffer, 1978; Nosofsky, 1990), generalize in learning (Pothos & Bailey, 2000), develop linguistic competence (Plunkett & Marchman, 1991), and solve reasoning problems (Kolodner, 1992; Tversky & Kahneman, 1983). Much of our thought appears to hinge on our ability to produce similarity intuitions (Goldstone & Son, 2005; Pothos, 2005). Yet, and despite intense effort, elucidating the principles that guide similarity judgments has proved elusive. In equal measure, a source of power of the similarity process and a challenge in understanding it is its flexibility: the same two objects can be judged as similar or dissimilar, depending on our purpose, perspective or context. Famously, Goodman's paradox is that any two objects can be considered arbitrarily similar, depending on which of their properties are considered (Goodman, 1972; cf. Barsalou, 1991). At first sight, similarity appears too unprincipled for formal study.

A way to harness similarity's multifaceted nature might be by assuming that its flexibility derives from contextual factors, so that, if such contextual factors can be understood, then we would likewise understand how similarity can appear to vary 'arbitrarily' (Goodman, 1972). Tversky (1977; Tversky & Gati, 1978) presented a highly cited example of the dependency of similarity on context. Consider a choice task of selecting the country most similar to Austria, amongst different choice sets. One choice set is {Sweden, Hungary, Poland} and another choice set is {Sweden, Hungary, Norway}. His results indicate that the similarity of Austria and Sweden, in a set with Hungary and Poland, was greater than the same similarity in a set with Norway and Hungary, a finding he called the diagnosticity effect. Apparently, the similarity between the same two countries, Austria and Sweden, depends on the choice set.

2. Empirical controversy

Diagnosticity has been hard to replicate. Glucksberg and Keysar (1990) observed a diagnosticity effect with choice sets which induced literal vs. metaphorical meanings. Medin and Kroll (unpublished, as cited in Medin, Goldstone, & Markman, 1995) distinguished between a choice task of selecting the option most similar to a target and a similarity task, corresponding to rank-ordering the similarity between all options and the target. They reported a diagnosticity effect with a similarity ranking task. Evers and Lakens (2014) examined evidence for diagnosticity involving both a choice task and a similarity ranking one. With the choice task, they reported a diagnosticity effect analogous to the one in Tversky (1977). However, with the similarity ranking task, there was a diagnosticity effect in only one of their studies (Study 1b); when they aggregated the results from all their ranking-task manipulations using meta-analytic methodology, they claimed the diagnosticity effect disappeared. Overall, there have been few replications of the diagnosticity effect and some anecdotal reports of replication failures (Evers and Lakens, 2014).

Diagnosticity is not the only possible contextual effect in similarity. Tversky (1977) found that in a choice set A, B, introducing an option X similar to A reduces preference for A, i.e., a reduction in preference for an option, when a distractor similar to the option is introduced. Are there reports for when introducing a distractor similar to an option *increases* preference (higher choice proportion or higher similarity between choice and target) for the option? In an important demonstration, Choplin and Hummel (2005) examined preference amongst one-dimensional, schematic stimuli. The choice set always included two options equally similar to a target and a decoy that would be just worse than one of the options. Employing both a choice task and similarity ratings, the introduction of the decoy made the option closest to it preferable – this is an effect which can be called *attraction*. Also, Trueblood et al. (2013) found both diagnosticity and attraction, employing two-dimensional schematic stimuli and a choice task. So, there is some limited evidence for an attraction effect in similarity judgments.

The attraction effect is opposite to the diagnosticity one, since in one case grouping an option with a distractor decreases preference for the option, but in the other it increases preference. For one-dimensional stimuli, diagnosticity and attraction cannot be manipulated independently, but it is convenient to retain separate labels. This leads to our proposal that there are two context effects in similarity, a diagnosticity effect and an attraction effect, in competition. Previous research is indicative of this possibility, but there has been no concurrent demonstration of the effects. If our proposal is supported, then any attempt to demonstrate one effect in isolation would be potentially undermined by the other. The apparent difficulty in replicating the diagnosticity effect can perhaps be traced to this reason.

Diagnosticity and attraction effects have been well-established in other areas. Note, there is a consistent definition of the attraction effect, regardless of the complexity of the stimuli (e.g., for multidimensional stimuli, the prerequisite for an attraction effect is the introduction of an option just inferior to a choice, on all dimensions). Maylor and Roberts (2007) reported both effects in episodic memory. Most demonstrations have been reported in decision making (where the diagnosticity effect is called a similarity effect, but we will retain the former term). In fact, an attraction effect in similarity judgments had been anticipated prior to Choplin and Hummel's (2005) work, by analogy with the decision literature (Dhar & Glazer, 1996; Medin et al., 1995). In consumer choice, with two dimensional stimuli, Huber, Payne, and Puto (1982) reported that introducing a decoy option increases preference for a slightly superior one (in relation to restaurant choice), which is an attraction effect. Attraction has also been demonstrated in political choice (Pan, O'Curry, & Pitts, 1995), selection of mates (Sedikides, Ariely, & Olsen, 1999), and assessing job candidates (Highhouse, 1996). Demonstrations of diagnosticity have been reported too, though not as frequently. Simonson (1989) observed diagnosticity (and attraction) in choice behavior involving consumer goods. Finally, there have been analogous reports in psychophysical tasks (Trueblood, Brown, & Heathcote, 2014; Tsetsos, Chater, & Usher, 2012). Concurrent demonstrations of attraction and diagnosticity in decision making (even with some controversy; Cataldo & Cohen, 2018a, 2018b; Frederick et al., 2014; Spektor et al., 2018; Yang & Lynn, 2014) further motivate corresponding tests in similarity.

3. Experiment 1

3.1 Participants

We recruited 210 participants via mTurk for a small fee. No exclusion criteria were adopted. Ethics was obtained approval from the Psychology Department, City, University of London.

3.2 Stimuli and procedure

We employed 17 spirals varying along a dimension of diameter size. The size of each spiral was given by the formula $S_n = S_0(1.1)^n$, with $S_0 = 7cm$ on our screen (the target) and $n = \{-8 \dots -1; 1 \dots 8\}$ (the 16 choice stimuli). According to Weber's law, $Sim(S_n, S_m) = Sim(S_{n+k}, S_{m+k})$, that is, similarity between spirals should depend only on *n* and the similarity between neighboring spirals should be constant, regardless of display screen resolution. In pilot testing we verified compliance with Weber's law (Electronic Supplementary Material Section, ESMS, 1).

The experiment involved a forced choice task of selecting the option most similar to a target spiral $T = S_0$, which was fixed across all trials. On each trial, three options were offered, two of which were equidistant to the target, denoted as S_A, S_B, and one which served as a distractor, S_C; so one option would always be smaller than the target and one larger. Each participant went through eight (Weber exponent for S_A varying from 1 to 8, for S_B from -1 to -8) x eight (Weber exponent for S_C varying from 1 to 8 or from -1 to -8) trials. That is, for each pair of S_A, S_B different trials were formed by varying the distractor S_C across all possible spiral sizes, except *C*=0. Between participants, S_C was either always smaller or always larger than the target. The arrangement of spirals within each trial was partly randomized; S_A, S_B randomly appeared on the left or right and S_C always appeared in the middle. Trial order was randomized. The three options for each trial were visible below the target until participants indicated their choice with a mouse click (Figure 1).

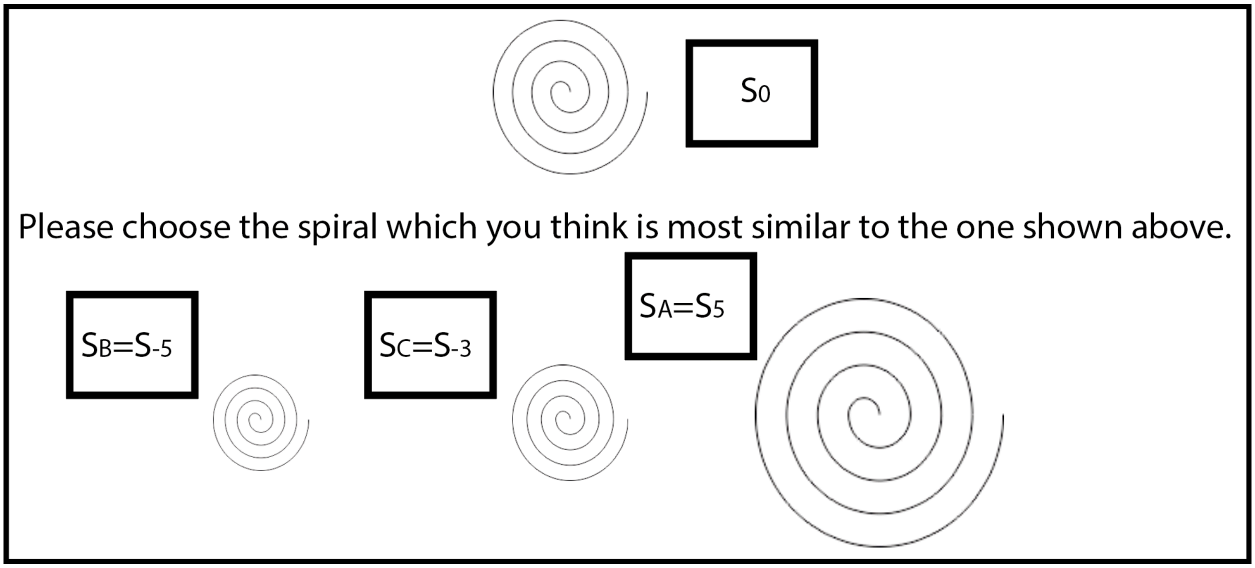

Figure 1. A typical trial in Experiment 1. The boxes were not shown to participants.

3.3 Empirical results and discussion

We present the data by collapsing across the between participant conditions, so that choice results are averaged for when the distractor item S_C is smaller or larger than the target. Figure 2 shows choice proportions for the three options, so that each graph shows trials with fixed $\{S_A, S_B\}$ and S_C varying relative to the target. For example, the graph for $Options\ at \pm 8$ shows average choice proportions for when S_C takes all possible values between 1 and ±8 steps away from the target. Regarding the competition between diagnosticity and attraction, the primary concern is whether the distractor is closer to fixed option S_A or S_B. Accordingly, we drop reference to S_A or S_B and present results for the fixed option Grouped with the distractor (blue curves) and the Ungrouped one (red curves). In the absence of contextual effects, the blue curves should coincide with the red ones. A diagnosticity effect is evident where the red curve is higher than the blue, green curves (preference for the Grouped option is reduced) and an attraction effect when the blue curve is higher than the red, green ones (preference for the Ungrouped option is increased).

The graphs for $Options\ at \pm 1,2,3,4$ most clearly show attraction and diagnosticity for the fixed options. When the distractor is closer to the target than S_A and S_B, the distractor dominates preference.

As the distractor shifts away from the target, it moves towards the Grouped option. When the Grouped option is close to the distractor, a diagnosticity effect emerges (higher preference for the Ungrouped option), but as soon as the Grouped option starts dominating the distractor, an attraction effect can be observed (higher preference for the Grouped option; Figure 3). Note, one advantage of employing one-dimensional stimuli is that there is clear dominance between options, translating to clear expectations for diagnosticity vs. attraction.

When the distractor is the same as the Grouped option, we do not observe equivalence of preference between the two. Possibly, this is because the distractor always appeared in the middle of the distal layout, introducing an asymmetry in presentation. Experiment 2 is a replication, avoiding this presentation bias, and increasing the trials per participant.

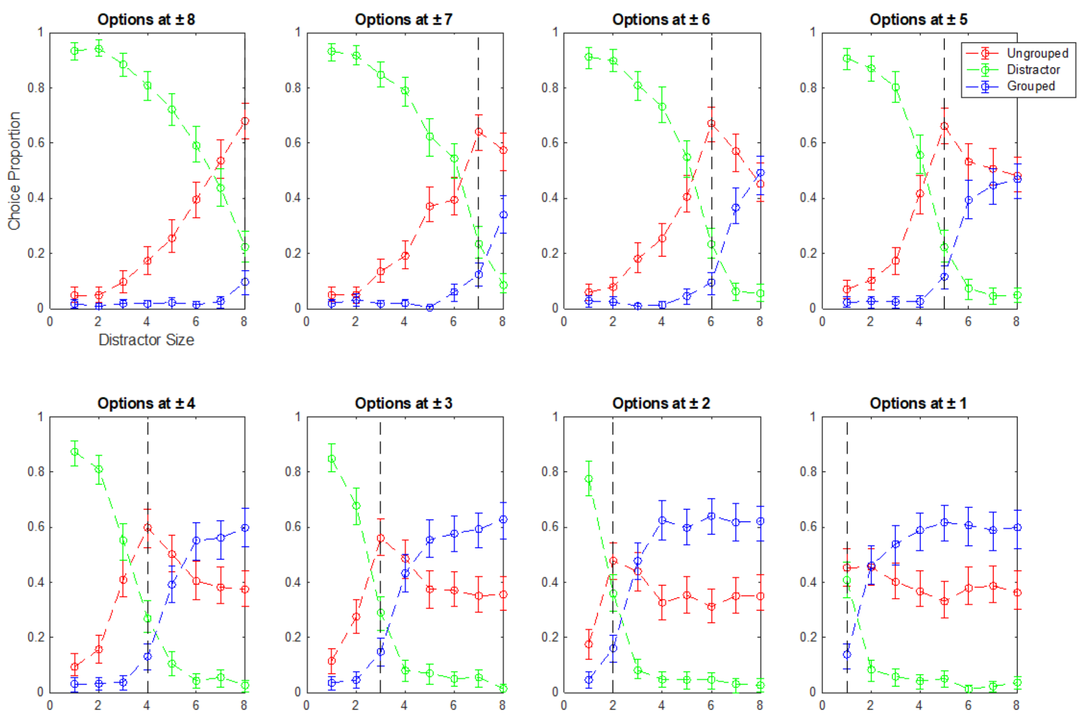

Figure 2. Experiment 1 choice proportions for the distractor (green curves), the item Grouped with the distractor (blue curves), and the Ungrouped item (red curves). The horizontal axis show Weber exponent and graphs are labelled by Weber exponent for the fixed options. Vertical lines indicate the fixed options. N=210 for each data point and error bars are 95% bootstrapped confidence intervals. Most panels show a transition from diagnosticity to attraction. For example, for Options at ± 4, when Weber exponent for S_C is close to 0, S_C is preferred (green curve > {red, blue curves}). As S_C gets closer to

Grouped, there is diagnosticity (red>blue for positions 2, 3, 4; red>green for 4). When S_C is clearly dominated by Grouped (6, 7, 8), there is attraction (blue>{red, green}).

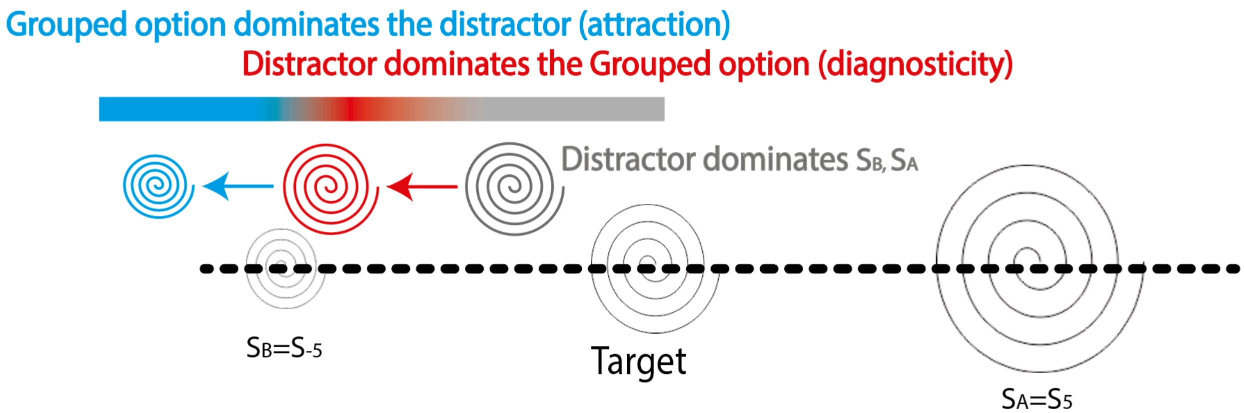

Figure 3. The transition from preference for the distractor (gray part of the bar), to diagnosticity (red part), to attraction (blue part) as the distractor dominates and is then dominated by the Grouped item (here S_B).

4. Experiment 2

4.1 Participants

We recruited 210 participants via mTurk for a small fee. No exclusion criteria were adopted.

4.2 Stimuli, procedure, results

We employed vertical lines varying in length, as in Choplin and Hummel (2005), to examine the replicability of the diagnosticity, attraction competition, with different stimuli. The length of each line was given by $S_n = S_0(1.05)^n$, with $S_0 = 4cm$ on the design computer. The similarity between lines should depend only on the power law exponent, i.e., $Sim(S_n, S_m) = Sim(S_{n+k}, S_{m+k})$. Trials were included to check consistency with Weber's law (ESMS 2).

Between participants, we adopted two sets of fixed options, with Weber exponents $A, B = \pm 4$ and $A, B = \pm 6$. The Weber exponents for the distractor were, in the former and latter cases respectively, $C = \{-8 \dots -2; 2 \dots 8\}$ and $C = \{-10 \dots -4; 4 \dots 10\}$. Each trial was a forced choice to determine which of the three options (S_A, S_B, S_C) was most similar to the target S_0, with the stimuli arranged as in Figure 1. Whether an option appeared in the left, middle, or right was random.

When the distractor is larger than the target, results cleanly reveal a transition from diagnosticity to attraction (right hand side of each panel in Figure 4). Additionally, there is preference neutrality when the distractor is identical to the Grouped option. When the distractor is smaller than the target (left hand side of each panel in Figure 7), there was no such pattern, but in this case stimulus control broke down (Weber's law was unconfirmed, ESMS 2).

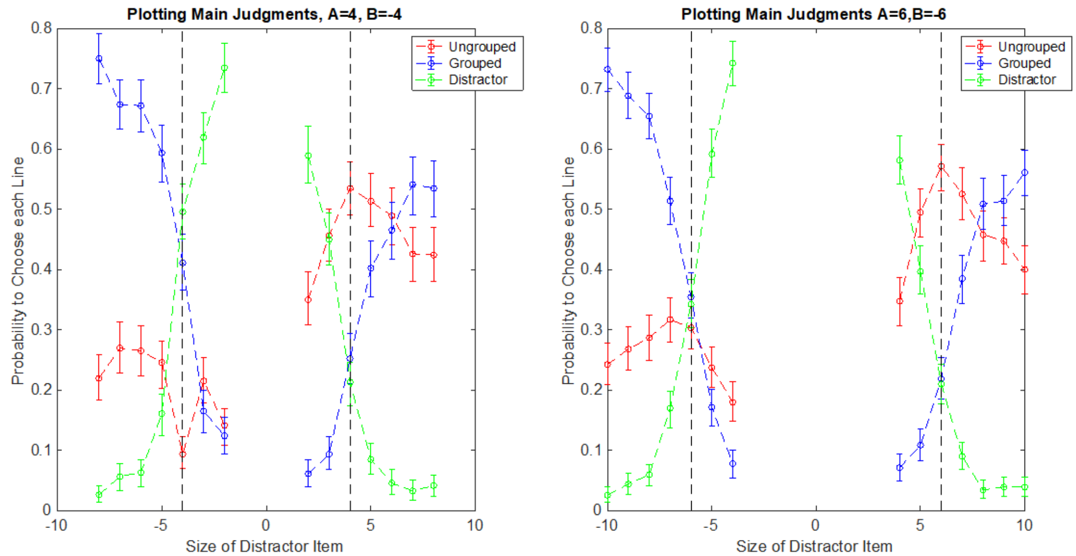

Figure 4. Experiment 2 choice proportions for the distractor (green curves), the item Grouped with the distractor (blue curves), and the Ungrouped item (red curves). The horizontal axis show Weber exponent and graphs are labelled by Weber exponent for the fixed options. Vertical lines indicate the fixed options. On the left hand side of each graph the Grouped item is the smaller one between S_A, S_B, and on the right hand side the Grouped item is the larger one. N=460 for each data point in the left graph and N=590 for each data point in the right graph, with each participant contributing five judgments; error bars are 95% bootstrapped confidence intervals. Diagnosticity requires red curve > {blue, green} curve and attraction blue > {red, green}.

5. Theory

Well-known similarity models have difficulty with the competition between diagnosticity, attraction. Tversky's (1977) proposal was that spontaneous grouping of the options shifts attention to different features, thus altering the similarity between the options. But Tversky's model predicts the same shift in similarity from grouping, regardless of whether the Grouped option dominates or is dominated by the distractor; so, for a particular grouping, it can only predict diagnosticity or attraction. The substitution effect is that preference for two grouped options is shared between them, and so reduced relative to an ungrouped option (Simonson, 1989); this clearly fails to account for attraction. In Krumhansl's (1978) distance-density model, similarity is a function of distance between stimuli, when represented in some geometric space, such that distance is greater when either stimulus is in a dense region of space. For a pair of stimuli, we can set model parameters to accommodate attraction or diagnosticity but not both. In Ashby and Perrin's (1988) General Recognition Theory (GRT), the effect of each stimulus is a distribution reflecting the varying perceptual effects of the stimulus. Responses for different stimuli correspond to different response regions in psychological space and the similarity between two stimuli depends on the overlap between the distribution of perceptual effects for the first stimulus and the response region for the second. The model accounts for diagnosticity, because when introducing a stimulus *C* close to *B*, the response region for *B* is reduced and so the similarity between *A* and *B* is reduced. We cannot see how this model can account for attraction.

The present results are problematic for some established decisions models. In multivariate decision field theory (MDFT), the available choices are compared in an attribute-specific way, with attribute selection at each time step determined stochastically (Roe, Busemeyer, & Townsend, 2001). For example, for two-dimensional options {*A*, *B*}, *A* may be superior to *B* on the first dimension, but inferior on the second. Regarding diagnosticity, momentary 'wins' would be split between grouped options, allowing the ungrouped one to be preferred. Regarding attraction, when a Grouped option clearly dominates another, a lateral inhibition mechanism provides additional support for it. However, these ideas require at least two-dimensional stimuli. In the Leaky Competing Accumulator (LCA) model, choices undergo momentary evaluations, until a decision criterion is reached (Usher & McClelland, 2004). The LCA employs an asymmetric, loss aversive value function (Tversky & Kahneman, 1991) to produce attraction, which restricts applicability to stimuli with hedonic properties. Finally, in the multivariate linear ballistic accumulator (MLBA; Trueblood, Heathcote, & Brown, 2014), diagnosticity requires an asymmetry in the decay of evidence from different types of comparisons, which is hard to motivate for meaningless stimuli.

We are not aware of simple, heuristic ideas which can cover attraction, diagnosticity competition for one-dimensional, meaningless stimuli. Instead, we outline a model, called the extended quantum similarity model (EQSM), based on previous work by Pothos, Busemeyer, and Trueblood (2013). Suppose we are interested in the similarity between stimulus *A* and a target stimulus *T*, denoted *Sim*(*A*,*T*). Pairwise similarity is modeled as thinking about *A* and then *T*, so that higher similarity means higher ease with which the transition can take place (cf. Hahn et al., 2003). The presence of contextual stimuli means that the basic thought process *A*→*T* is supplemented with additional trains of thought involving the contextual stimuli, which can help *focus or distract from* the similarity judgment. An example of the former is $Sim(Hyenas, Lions; Elephants)$. The contextual item is Elephants and is separated from the compared items, Hyenas and Lions, by a ';'. The target is Lions. Here, the contextual item helps activate features that are shared between the Hyenas and Lions (Tversky, 1977). Such focusing trains of thought can be denoted as *Context*→*A*→*T*. An example of a distracting train of thought is when the contextual item interferes with the comparison we are interested in, so that we have *A*→*Context*→*T*. Note, we assume that any train of thought can be operationalized as ease of thinking of a series of stimuli, one by one.

Our proposal is therefore:

$$Sim(A,T;Context1,Context2\ldots) = Sim(A,T) + \alpha \cdot Focusing - \beta \cdot Distracting \ldots\ldots \text{Equation (1)}$$

where Focusing refers to all the focusing trains of thought and analogously for Distracting, and α, β are free parameters. Focusing trains of thought contribute to a similarity judgment, because they are consistent with the requirement that the similarity between *A* and *T* involves a transition from *A* to *T*. For distracting trains of thought, there is no longer a direct transition from *A* to *T*. Since such trains of thought fail to match intention, we assume they lower the similarity between *A* and *T*.

This proposal is easy to formalize using quantum probability theory (QPT), the probability rules from quantum mechanics without the physics (Busemeyer & Bruza, 2011; Pothos & Busemeyer, 2013), and the resulting model is EQSM. In QPT, the mental state is a normalized vector and different thoughts are subspaces in a multidimensional space. Then, a train of thought can be operationalized as a sequence of 'projections' of the mental state onto the subspaces corresponding to the relevant stimuli. Using this idea and a restriction to one-dimensional subspaces, the EQSM Equation (1) becomes:

$$Sim(A,T;Context1,Context2) = Sim(A,T) + \alpha[Sim(Context1,A)Sim(A,T) + Sim(Context2,A)Sim(A,T)] - \beta[Sim(A,Context1)Sim(Context1,T) + Sim(A,Context2)Sim(Context2,T)] \ldots\ldots \text{Equation (2)}$$

Equation (2) allows modeling context using just pairwise similarities. Overall, a prediction of diagnosticity vs. attraction will depend both on the parameters and stimulus similarities. The model can describe results from Experiments 1, 2 well. Deviance Information Criterion difference between the EQSM and a matched model with no contextual influences was approximately 130 for Experiment 1 and 80 for Experiment 2 (ESMS 3). Moreover, as argued, it is unclear how established similarity and decision models can capture the competition between diagnosticity, attraction.

6. Concluding comments

The flexibility of similarity may be partly due to context effects. Even though such context effects have been well-known, and analogous effects have been observed in decision-making, we identified important limitations: in similarity, researchers have been surprisingly restricted to diagnosticity (Choplin & Hummel's, 2005, pioneering work is an exception) and the replicability of diagnosticity has been a major challenge; in decision-making, research has typically involved complex stimuli. A lack of demonstration of concurrent attraction, diagnosticity, for one-dimensional, meaningless stimuli is a shortcoming of both literatures. We outlined a similarity model to cover our data, based on operationalising contextual influences as focussing and distracting thought processes. This project originated as a way to provide more ambitious datasets for the original quantum similarity model (Pothos et al., 2013).

In similarity, one-dimensional stimuli offer the cleanest demonstration of the attraction, diagnosticity competition. Extending to multi-dimensional stimuli is complicated because differences in stimulus complexity may lead to asymmetries in judgment (Tversky, 1977). This is a worthwhile objective for future research.

Acknowledgements

EMP was supported by Office of Naval Research Global grant N62909-19-1-2000. JST was supported by NSF grant 1846764.

References

Aguilar, C. M., & Medin, D. L. (1999). Asymmetries of comparison. Psychonomic Bulletin & Review, 6, 328-337.

Ashby, G. F. & Perrin, N. A. (1988). Towards a Unified Theory of Similarity and Recognition. Psychological Review, 95, 124-150.

Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), The psychology of learning and motivation, Vol. 27, pp. 1-64. New York: Academic Press.

Busemeyer, J. R. & Bruza, P. (2011). Quantum models of cognition and decision making. Cambridge University Press: Cambridge, UK.

Cataldo, A. M., & Cohen, A. L. (2018a). Reversing the similarity effect: The effect of presentation format. Cognition, 175, 141-156.

Cataldo, A. M., & Cohen, A. L. (2018b). The comparison process as an account of variation in the attraction, compromise, and similarity effects. Psychonomic bulletin & review, 1-9.

Choplin, J. M. & Hummel, J. E. (2005). Comparison-induced decoy effects. Memory & Cognition, 33, 332-343.

Dhar, R., & Glazer, R. (1996). Similarity in context: Cognitive representation and violation of preference and perceptual invariance in consumer choice. Organizational Behavior and Human Decision Processes, 67, 280–293.

Evans, N. J., Holmes, W. R., & Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. Psychonomic bulletin & review, 1-33.

Evers, E. R. K. & Lakens, D. (2014). Revisiting Tversky's diagnosticity principle. Frontiers in Psychology, Article 875.

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. Journal of Marketing Research, 51, 487-507.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. Psychological Review, 97, 3-18.

Goldstone, R. L, & Son, J. (2005). Similarity. In K. Holyoak & R. Morrison (Eds.). Cambridge Handbook of Thinking and Reasoning. Cambridge: Cambridge University Press. (pp. 13-36).

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman, Problems and projects (pp. 437-447). Indianapolis: Bobbs-Merrill.

Hahn, U., Chater, N., & Richardson, L. B. (2003). Similarity as transformation. Cognition, 87, 1-32.

Highhouse, S. (1996). Context-dependent selection: The effects of decoy and phantom job candidates. Organizational Behavior and Human Decision Processes, 65, 68–76.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. Journal of Consumer Research, 9, 90-98.

Kolodner, J. L. (1992). An introduction to Case-Based Reasoning. Artificial Intelligence Review, 6, 3-34.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. Psychological Review, 85, 445-463.

Medin, D. L., & Kroll, J. (1994). Studies of diagnosticity involving attributes and relations. Manuscript in preparation.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

Medin, D. L., Pineda, K., & Markman, A. M. (1994, May). Parallels between similarity making and decision making. Paper presented at the 66th annual meeting of the Midwestern Psychological Association, Chicago, IL.

Medin, D. L., Goldstone, R. L., & Markman, A. B. (1995). Comparison and choice: relations between similarity processes and decision processes. Psychonomic Bulletin & Review 2, 1–19.

Nosofsky, R. M. (1990). Relations between exemplar-similarity and likelihood models of classification. Journal of Mathematical Psychology, 34, 393-418.

Pan, Y., O'Curry, S., & Pitts, R. (1995). The attraction effect and political choice in two elections. Journal of Consumer Psychology, 4, 85–101.

Plummer, M., et al. (2003). Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing (Vol. 124, p. 125).

Plunkett, K. & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. Cognition, 38, 43-102.

Pothos, E. M. (2005). The rules versus similarity distinction. Behavioral & Brain Sciences, 28, 1-49.

Pothos, E. M. & Bailey, T. M. (2000). The Importance of Similarity in Artificial Grammar Learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 847-862.

Pothos, E.M. & Busemeyer, J.R. (2013). Can quantum probability provide a new direction for cognitive modeling? Behavioral & Brain Sciences, 36, 255-327.

Pothos, E. M., Busemeyer, J. R., & Trueblood, J. S. (2013). A quantum geometric model of similarity. Psychological Review, 120, 679-696.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. Psychological Review, 108, 370-392.

Sedikides, C., Ariely, D., & Olsen, N. (1999). Contextual and procedural determinants of partner selection: Of asymmetric dominance and prominence. Social Cognition, 17, 118-139.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. Journal of Consumer Research, 16, 158–174.

Spektor, M. S., Kellen, D., & Hotaling, J. M. (2018). When the good looks bad: An experimental exploration of the repulsion effect. Psychological science, 29, 1309-1320.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64, 583–639.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. Psychological Science, 24, 901–908.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. Psychological Review, 121, 179-205.

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multi attribute choice. Psychological Review, 117, 1275-1293.

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. Proceedings of the National Academy of Sciences of the United States of America, 109, 9659–9664.

Tversky, A. (1977). Features of Similarity. Psychological Review, 84, 327-352.

Tversky, A. & Gati, I. (1978) Studies of similarity. In: Cognition and categorization, ed. Rosch, E. & Lloyd, B. B., pp. 79–98. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. Quarterly Journal of Economics, 107, 1039-1061.

Tversky, A., & Simonson, I. (1993). Context dependent preferences. Management Science, 39, 1179-1189.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. Psychological Review, 111, 757-769.

Yang, S., & Lynn, M. (2014). More evidence challenging the robustness and usefulness of the attraction effect. Journal of Marketing Research, 51(4), 508-513.

Electronic Supplementary Material Section 1. Experiment 1 compliance with Weber's law

We report a pilot study showing compliance with Weber's law, in Experiment 1.

A separate participant sample was recruited as for the main study (N=100) and tested with a pairwise similarity task, involving all spiral pairs. Responses on a 1-9 scale were linearly transformed within participants to ensure consistent use of anchor points. For *n-m*>0, plotting $\log S_{n-m}$ (i.e., the logarithm of the ratings) against *n-m* (the Weber exponents for compared stimuli S_n, S_m) should produce a straight line and indeed this was the observed result (R^2>0.99). For *n-m*<0 we obtained a similar result, but because of a computer error the Weber fraction was slightly different from the intended one (i.e., the Weber fraction for stimuli smaller than S_0 was unintentionally slightly different to that for stimuli larger than S_0). Either way, there was good evidence for Weber's law validity across the stimulus range and discriminability of stimuli. Figure S1 shows the relevant plot.

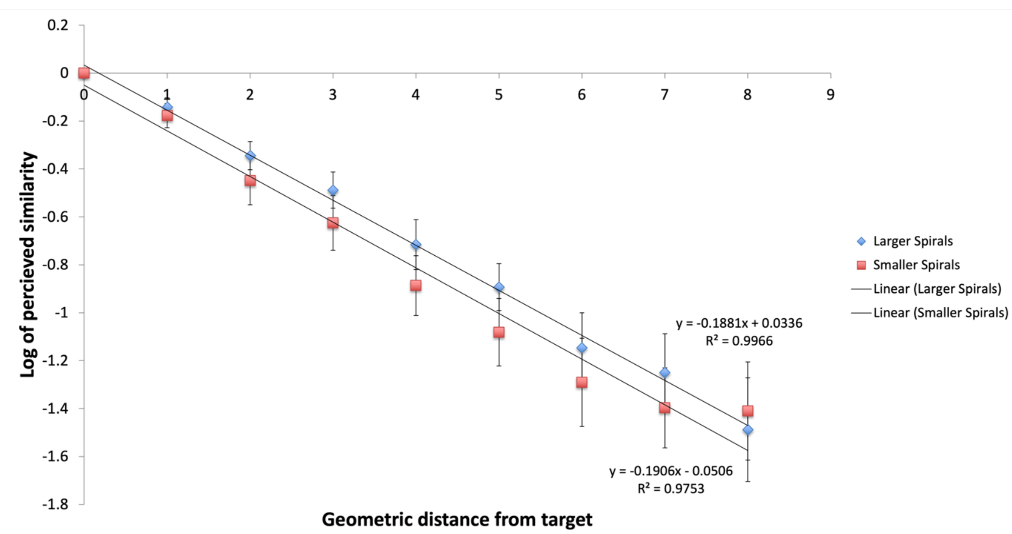

Figure S1. Plotting the log of perceived similarity vs. geometric distance from target, for the stimuli in Experiment 1.

<u>Electronic Supplementary Material Section 2. Experiment 2 compliance with Weber's law</u>

We report the results from trials included in Experiment 2 to test compliance with Weber's law.

To check the validity of Weber's law for the stimuli in Experiment 2, we included some forced choice trials with only two options, S_X, S_Y. In both between participant conditions, the target for these trials was as in the main experiment. In the $A, -A = \pm 4$ between participants condition, for the first option, $X = \pm 4$, while for the second option, $Y = \{-4.2 \ldots - 3.4; 3.4 \ldots 4.2\}$. This arrangement produces four combinations, e.g., when $X = 4$ and $Y = \{-4.2 \ldots - 3.4\}$ etc. Notice that preference between S_X, S_Y should be neutral when $Y = \pm 4$, smoothly increase with decreasing size for S_Y, and smoothly decrease with increasing size for S_Y. For the $A, -A = \pm 6$ between participants condition, analogous stimuli were employed. The results are shown in Figure S2 and the key conclusion is that there is a uniform higher preference for smaller vs. larger stimuli and in particular $Sim(S_{X=-4}, T) > Sim(S_{Y=4}, T)$ and $Sim(S_{X=-6}, T) > Sim(S_{Y=6}, T)$. This bias plausibly arises from non-applicability of Weber's law for smaller stimuli.

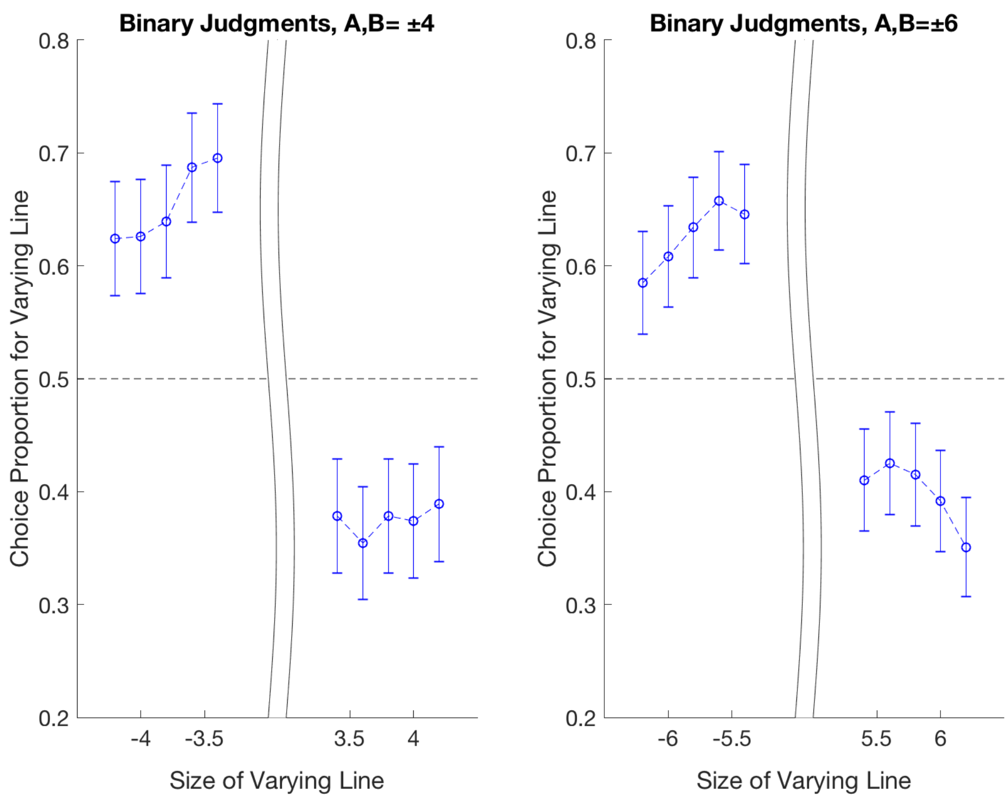

Figure S2. Choice proportions in the two option trials, employed to investigate Weber's law for the present stimuli. For the left-hand plot one line was fixed at ±4 and the other was allowed to vary around ∓4. The lines indicate preference for the varying length line, as it took lengths from ±4.2 to ±3.4. The right-hand plot is similar, but with the fixed lines at ±6. There are two general features: first, preference for the varying line (analogous to the distractor in the main experiments) increases as it moves closer to 0 (the length of the target), as expected. Second, there is an overall preference for shorter over longer lines, indicated by the fact the choice proportions do not approach 0.5 even when the varying line is as ±4 or ±6.

Electronic Supplementary Material Section 3. Outline of the model and fits

We present some details of the Extended Quantum Similarity Model (EQSM), which is based on Quantum Probability Theory (QPT). QPT may be unfamiliar to readers. However, especially in similarity, QPT is a natural extension of the standard geometric models of similarity. In standard geometric models of similarity, each stimulus is a point is a multidimensional space and similarity is a function of distance (usually an exponentially decaying function). In QPT, each stimulus is a subspace. Similarity is computed as projection of a state vector (representing the mental state) from the subspace for the first stimulus to the subspace for the second one (Figure S3). That is, similarity depends on the angle(s) between subspaces.

Subspaces can be at different angles, so this approach embodies a straightforward way to operationalize ease of thinking of one stimulus given another, as the ease of transition across two subspaces. More specifically, a main idea is that sequences of thoughts are sequences of projections across subspaces; these sequences of projections embody the contextuality necessary for attraction or diagnosticity. The main part of our proposal is that context arises from focusing or distracting trains of thought – QPT provides a particularly easy way to implement this idea.

The model details (just below) are mostly linear algebra. We proceed with a summary of the model and fits. Readers interested in QPT are further referred to Pothos et al. (2013), who presented the original QPT similarity model, and Busemeyer and Bruza (2011), who offer a more extensive tutorial.

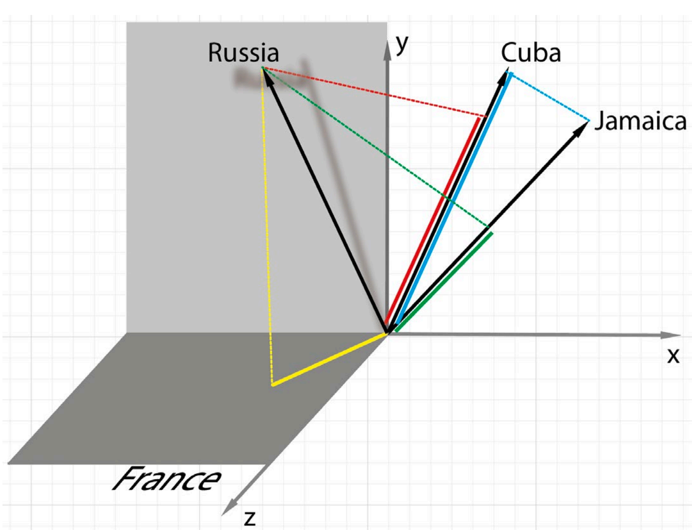

Figure S3. An illustration of projection. Jamaica, Cuba, and Russia are represented as one-dimensional subspaces and France as a two-dimensional one. The projection from Jamaica to Cuba is large (blue solid line), because the two subspaces are 'close' to each other; this is less so for Cuba and Russia (solid red line, projecting from Russia to Cuba) or for Jamaica and Russia (solid green line, projecting Russia to Jamaica). We can also project to a two-dimensional subspace, which is like 'casting a shadow' of the state vector onto the subspace. In the Russia, France example, we cast a shadow of a vector along Russia onto France. Projecting from France Russia would entail projecting a (normalized) vector within France to Russia.

The mental state is represented as $\rho = \sum w_i |\psi_i\rangle\langle\psi_i|$, where w_i's are classical probabilities that the actual state vector is $|\psi_i\rangle$. We employ Dirac notation, so that $|\psi_i\rangle$ denotes a column vector and $\langle\psi_i|$ its complex conjugate transpose. Prior to a similarity judgment between some stimuli the mental state should reflect uninformativeness between the stimuli, i.e., $\rho = I$, where I is the identity matrix (this is equivalent to a uniform prior in Bayesian statistics). A measure of the overlap between ρ and a subspace *A* (of any dimensionality) is the probability that ρ is about *A*, $Prob(\rho = A) = Tr(\rho \cdot P_A)$, where P_A is the projection operator to subspace A and *Tr* is the trace operator, which sums the diagonal elements of a matrix. This probability rule is a generalization of the projection ideas above.

The basic definition of similarity is $Sim(A, T) = Tr(P_T P_A \rho P_A)$, and it arises from a simple train of thought from the reference item (*A*) to the target one (*T*). Similarity in the presence of contextual item *B* involves *focusing* trains of thought, $\rho \to B \to A \to T$, and *distracting* trains of thought, $\rho \to A \to B \to T$, where e.g. the former is given by $Tr(P_T P_A P_B \rho P_B P_A)$. Then, the EQSM for the similarity between a target *T* and a choice element *A*, in the presence of contextual stimuli *B*, *C* is given by:

$$Sim(A, T; B, C) = Tr(P_T P_A \rho P_A) + \alpha[Tr(P_T P_A P_B \rho P_B P_A) + Tr(P_T P_A P_C \rho P_C P_A)] - \beta[Tr(P_T P_B P_A \rho P_A P_B) + Tr(P_T P_C P_A \rho P_A P_C)]$$ ……………………………………………………………….Equation 1A

The parameters α, β concern the weight of focusing trains of thought relative to distracting ones. For one-dimensional stimuli, we can readily specify the various projection operators, e.g., $P_A = |A\rangle\langle A|$. Then, $Tr(P_T P_A P_B \rho P_B P_A) = \langle T|A\rangle\langle A|B\rangle\langle B|\rho|B\rangle\langle B|A\rangle\langle A|T\rangle = \langle B|\rho|B\rangle |\langle A|B\rangle|^2 |\langle T|A\rangle|^2 = Sim(B, A) \cdot Sim(A, T)$. These simplifications lead to Equation (2) in main text.

Regarding model fits, because the stimuli are one-dimensional, we can employ the simplified EQSM (Equation 2 in main text). A standard way to compute pairwise similarities from distances in a vector space is Shepard's (1987) law, $Sim(A, B) = e^{-\gamma * d(A,B)^2}$, where $d(A, B)$ is the distance between stimuli expressed in Weber fractions, and γ is a positive constant, which determines the rate of change of similarity with distance (often referred to as the sensitivity parameter). Note, we are employing the Gaussian version of Shepard's law, because in both experiments the stimuli can be assumed confusable (Nosofsky, 1992). Similarity differences must be converted to choice probabilities, and this is achieved through the SoftMax rule, which converts an unnormalized vector into a probability distribution (e.g., Trueblood et al., 2017). Specifically, the proportion of choosing *A* as the option most similar to *T*, with other options *B*, *C* is given by $SoftMax(A) = \frac{\exp\left(\frac{Sim(A,T;B,C)}{\tau}\right)}{\exp\left(\frac{Sim(A,T;B,C)}{\tau}\right)+\exp\left(\frac{Sim(B,T;A,C)}{\tau}\right)+\exp\left(\frac{Sim(C,T;B,A)}{\tau}\right)}$, where $\tau > 0$ is a temperature parameter. Higher values of τ produce more deterministic responding. Finally, we assume that judgments are drawn from a Dirichlet distribution with parameters $\{\lambda * SoftMax(A), \lambda * SoftMax(B), \lambda * SoftMax(C)\}$, where $\lambda > 0$ is a shape parameter. Overall, this application of EQSM has five parameters, three associated with EQSM processes (α, β, γ) and two relating to the translation of similarities into choice proportions (τ, λ). The critical EQSM parameters are α, β, which determine the balance of contributing vs. distracting thought sequences.

The quantum model was fitted to experimental data using MCMC Gibbs sampling in JAGS (Plummer et al., 2003).

We first consider fits to the Experiment 1 data. To summarize, the similarity of each option to the target, given the presence of different context elements, is computed using the various e.g. $Sim(A, T; B, C)$ quantities; a choice proportion is computed using these quantities. Each contextual similarity $Sim(A, T; B, C)$ can be computed using pairwise similarities (Equation 2 in main text), which can be estimated using Weber's law e.g. $Sim(A, B) = e^{-\gamma * d(A,B)^2}$, where the distances correspond to the experimenter-defined representation of the stimuli. All these details apply to Experiment 2 as well, but for the differences concerning Weber's law, outlined below.

We fitted the average (across participants) choice proportions for each combination of fixed option and distractor sizes, but without collapsing across the between participant conditions when the distractor was smaller and larger than the target. We employed uninformative priors for all parameters, namely $\alpha \sim Uniform(0,5)$, $\beta \sim Uniform(0,5)$, $\gamma \sim Uniform(0,.1)$, $\lambda \sim Uniform(1,200)$, $\tau \sim Normal(.1,100)$ truncated to lie in $[0,1]$. We employed three MCMC chains, with 50,000 samples and a burn-in of 5,000 samples. Convergence was assessed with the $\hat{R}$ statistic and all chains displayed good convergence.

EQSM fit was very good. Any misfitting does not appear systematic and the model can correctly capture the transition from diagnosticity to attraction, which occurs as the distractor changes from being dominant to being dominated in relation to the Grouped option. Figure S4 provides an example plot for an illustrative combination of stimulus sizes. To formally assess EQSM performance, we specified a null model corresponding to setting $\alpha = \beta = 0$, that is, the null model attempts to fit data on the basis of pairwise similarity information, but without the benefit of contextual effects. Using the Deviance Information Criterion (DIC; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002), which is a statistic for comparing non-nested models taking into account model complexity, the DIC difference between the EQSM and the null model was approximately 130, which is strong evidence for the model over the null. We considered parameter posterior distributions and there is no indication of problematic behavior.

Regarding the Experiment 2 data, the results concerning Weber's law indicate a preference for the shorter over the longer lines. A simple way to account for this is by assuming that the similarity scaling through Shepard's law involves different sensitivity parameters for smaller vs. larger stimuli, relative to the target, which we denote as γ^- vs. γ^+ respectively. Uninformative priors were employed for these parameters, $\gamma^+ \sim Uniform(0, .1)$ and $\gamma^- \sim Uniform(0, .1)$, while the priors for the other parameters were as for the Experiment 1 model fits. Number of chains, burn-in, and other details as above.

We first fitted the EQSM to choice proportions computed by averaging across participants for each unique combination of stimulus sizes, as in Experiment 1. The results for $A, B = \pm 6$ are shown in Figure S5 (results for $A, B = \pm 4$ are similar). The model closely follows behavioral patterns in all cases, except for part of the stimulus range for the two-option trials when $A, B = \pm 4$. The EQSM again reproduces the transition from attraction, when the distractor is dominated by the Grouped option, to diagnosticity, when the distractor is very similar to the Grouped option, to no contextuality as the grouping between the distractor and the Grouped option breaks down. The DIC difference between the EQSM fit and the fit of a matched model but without contextuality ($\alpha = \beta = 0$, as above) was approximately 80, confirming the substantial advantage of the former. Note, this shows that the modeling approach can handle non-consistency with Weber's law. That is, as noted, the data we have from Experiment 2 is not consistent with Weber's law, with participants showing an overall preference for shorter lines. This upsets the symmetry of the stimuli and leads to a breakdown of the expected context effects when the distractor is smaller than the target (left hand side of the plots in Figure 4). However, the model is still able to account for the data by incorporating this observed violation of Weber's law, and the model fits (Figure S5) are equally good for distractor items smaller or larger than the target.

In Experiment 2, there were multiple judgments per *A*, *B*, *C* combination per participant. Accordingly, we can examine whether the crossover between attraction and diagnosticity is reflected in individual behavior or whether there are participants who predominantly show attraction or diagnosticity, but not both. We adopted a behavioral measure for grouping participants into homogeneous groups, based on whether they demonstrated attraction, diagnosticity, or both; participants were approximately evenly divided amongst these three groups (note, in Experiment 2 the two effects were concurrently present only when the distractor was larger than the target, so participant classification was based on these trials). The group sizes we observed were as follows: for when $A, B = \pm 4$, attraction only N=32, diagnosticity only N=34, and both effects N=26; for when $A, B = \pm 6$, attraction only N=38, diagnosticity only N=43, and both effects N=37. This classification establishes that a crossover pattern between attraction and diagnosticity is not an artifact of averaging.

Fitting the full and the null models to each group separately still showed an advantage for the former, with a DIC~20 for the groups displaying only attraction or diagnosticity and DIC~30 for the group displaying both effects. Thus, a contextual similarity model is needed for adequate description of the results, even when a clear pattern of contextuality is undermined by inconsistent stimulus perception. The conclusions for the $A, B = \pm 6$ trials are essentially identical (DIC values were larger).

A final issue is whether the success of the EQSM might be attributable to flexibility. We offer two indicative analyses that this is not the case. First, we created an artificial dataset, for which the pattern of diagnosticity vs. attraction is opposite to the one observed: when the distractor just dominates the target, there is increased preference for the grouped option (attraction, instead of diagnosticity) and when the grouped option just dominates the distractor, there is decreased preference for the grouped option (diagnosticity, instead of attraction). Figure S6a shows the artificial dataset and Figure S6b the fits. Clearly, the EQSM is no longer able to capture the ‘observed’ crossover between diagnosticity and attraction. We note that the root-mean-square error for the EQSM for the Experiment 2 dataset (for the ± 4 options) was $6 \cdot 10^{-2}$, but for the artificial dataset 50% more at $9 \cdot 10^{-2}$. Second, we created an artificial dataset, analogous to that for ± 4 in Experiment 2, but without context effects (to keep things interesting, we included the preference for shorter lines). Here, the null model without contextual effects produced a superior fit (DIC difference of 4).

In sum, it is easy to create a plausible artificial dataset for which the EQSM performs poorly. This shows that the model cannot fit any arbitrary data pattern equally well. Moreover, for a dataset with no contextual influences, a null model, matched to the EQSM but without contextual influences, outperforms the EQSM.

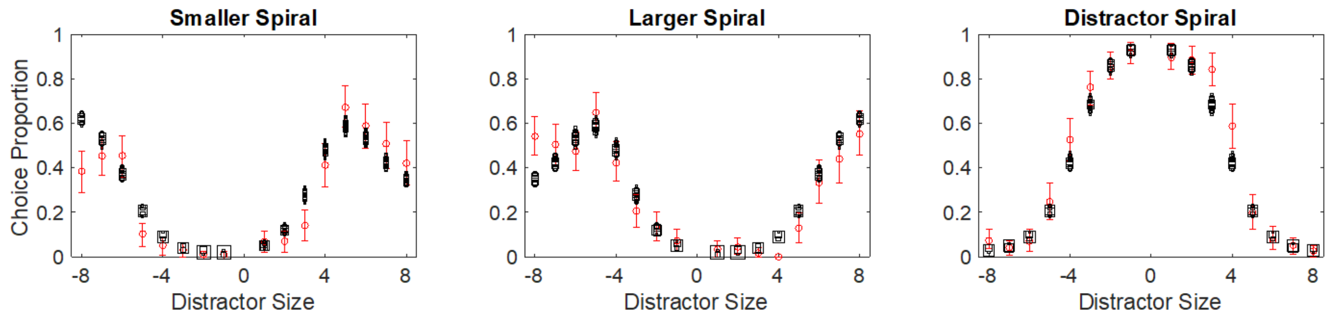

Figure S4. An example plot of posterior predictive distributions (black squares) and choice proportions (red dots) for EQSM fits for the $A, B = \pm 5$, in Experiment 1. The left panel corresponds to the smaller spiral (S_A), the middle to the larger one (S_B), and the right to the distractor one (S_C). Posterior distributions are represented by histograms, with the area of the square proportional to the posterior probability. Error bars for the data are bootstrapped 95% confidence intervals.

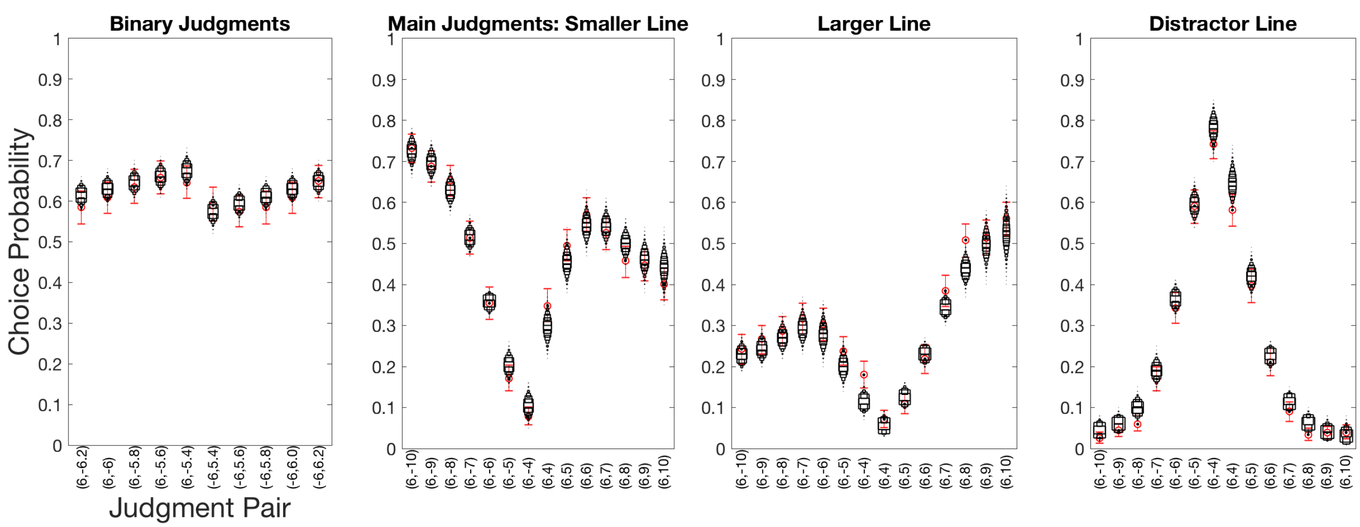

Figure S5. Posterior predicted distributions (black squares) and choice proportions (red dots) for EQSM fits for the $A, B = \pm 6$ trials Experiment 2. Posterior distributions are represented by histograms, with the area of the square proportional to the posterior probability. Error bars for data are bootstrapped 95% confidence intervals. In the leftmost panel, we see the fits to the two-option trials, which were employed for testing Weber's law validity for these stimuli. The rest of the panels show EQSM fits for the three-option trials. Choice proportions for options S_B (small line option), S_A (large line option), and

S_C (distractor) are shown in different panels moving from the left to the right. The horizontal axis shows the size of A (=-B) and the distractor.

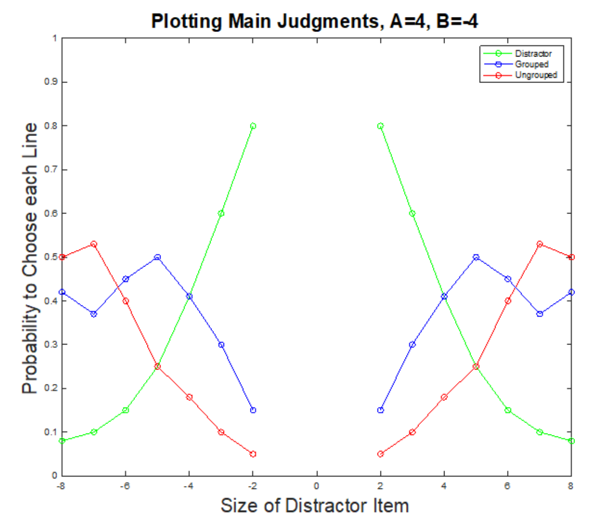

Figure S6a

Figure S6b

Figure S6. On the left (S6a), we show an artificial dataset constructed to reflect a cross-over pattern of attraction vs. diagnosticity opposite to that observed for the same stimuli in Experiment 2 (constructed without the bias for shorter lines, for simplicity). The blue, red, and green curves show choice preference for the grouped, ungrouped, and distractor options, respectively. On the left (S6b), we show best EQSM fit.

<u>References not in main text</u>

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. Annual Review of Psychology, 43, 25-53.

Plummer, M., et al. (2003). Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing (Vol. 124, p. 125).

Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. Science, 237, 1317-1323.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64, 583–639.