Chapter 1

Similarity Judgments: From Classical to Complex Vector Psychological Spaces

Albert Barque Duran, Emmanuel M. Pothos, James M. Yearsley, James A. Hampton

Department of Psychology, City University London, London, UK

Jerome R. Busemeyer

Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

Jennifer S. Trueblood

Department of Cognitive Science, University of California, Irvine, CA, USA

This chapter reviews progress with applications of quantum theory in understanding human similarity judgments. We first motivate and subsequently describe the quantum similarity model (QSM), which was proposed by Pothos, Busemeyer and Trueblood [2013], primarily as a way to cover the empirical findings reported in Tversky [1977]. We then show how the QSM encompasses Tversky's [1977] results, specifically in relation to violations of symmetry, violations of the triangle inequality and the diagnosticity effect. We next consider a list of challenges of the QSM and open issues for further research.

1. Background and motivations for a new model

Similarity judgments play a central role in many areas of psychology [e.g. Goldstone, 1994; Medin, Goldstone & Gentner, 1993; Pothos, 2005, Sloman & Rips, 1998]. Consequently, they have received much attention [e.g. Goodman, 1972], especially in relation to Tversky's [1977] findings, which have been a major focus of subsequent theoretical work on similarity judgments.

One traditional way to understand similarity uses a geometric approach, whereby similarity is assumed to correspond to a function of the distance between concepts in a psychological space. According to this approach, stimuli or concepts are represented as points in a multidimensional psychological space, with similarity being a decreasing function of distance in that space. The origin of the debate, criticism and the several attempts to empirically refute this approach all relate to the fact that similarity measures based on distance must obey various properties, called the metric axioms, that all distances (and simple related measures) are subject to. The most famous demonstration that human similarity judgments are inconsistent with these properties is due to Tversky [1977]. The importance and the impact of Tversky's paper come from the fact that his findings questioned the fundamental properties of any model of similarity based on distance in psychological space. Specifically, Tversky's approach was to provide empirical tests of the metric axioms, regardless of the specifics of the similarity approach. Showing, as he did, that the metric axioms are inconsistent with human similarity judgments, he concluded that human similarity judgments cannot be modelled with any distance-based approach.

Specifically, Tversky [1977] reported violations of (1) minimality: identical objects are not always judged to be maximally similar; (2) symmetry: the similarity of A to B can be different from that of B to A; (3) the triangle inequality: the distance between two points cannot exceed the sum of their distances to any third point; (4) a diagnosticity effect: the similarity between the same two objects can be affected by which other objects are present. In the next four sections we elaborate on all these findings and we consider some notable previous theoretical efforts to account for Tversky's [1977] challenges. Note we do not consider minimality, since most models (including the QSM) can become consistent with violations of minimality through some process of noise in how representations are specified and compared.

1.1. *Asymmetries*

If similarity is determined by distance, then how could it be the case that the similarity between two objects depends on the order in which the objects are considered? Directionality can arise from the fact that the relevant stimuli are not (always) simultaneously presented. For example, the temporal ordering of the stimuli can impose directionality structure in the similarity comparison. Alternatively, directionality can be conveyed in a syntactical way, e.g., if an observer is asked to evaluate sentences like "A is similar to B". Whenever this happens, there is a potential for asymmetry. This can be readily seen in the kind of task Tversky [1977] employed to explore putative violations of symmetry. He asked participants to choose which they preferred between these two statements: "North Korea is similar to Red China" and "Red China is similar to North Korea" (for simplicity we will use only Korea and China). Most participants preferred the former to the latter statement (this demonstration involved several other pairs of counties and was generalized to other kinds of stimuli). This result implied that the similarity of Korea to China (expressed as sim(Korea, China)) is higher that of China to Korea (expressed as Sim(China, Korea)), and thus revealed a violation of symmetry in similarity. Tversky's interpretation about why such asymmetries arise related to differences between the two stimuli in the extent of featural knowledge combined with differential weight given to the features specific to each concept (the parameters α and b in his model, see below). But, asymmetries in similarity judgments can also arise in other ways: Polk et al. [2002] proposed that they can also be the result of differences in the frequency of occurrence of one of the compared stimuli (a higher similarity was observed when comparing a low frequency stimulus with a high frequency one). Even before Tversky's [1977] work, Rosch [1975] had proposed similarity asymmetries can arise when a less prototypical stimulus is compared to a more prototypical one.

Asymmetries are difficult to reconcile with the idea of similarity-as-distance. Some kind of mechanism that can produce asymmetries, in some circumstances, in a more natural way is clearly desirable. We will see below that a quantum approach provides such a mechanism.

1.2. *The triangle inequality*

Tversky [1977] also considered how similarity judgments can lead to violations of the triangle inequality, another one of the metric axioms. In his paper, he states that (p.329) "the perceived distance of Jamaica to Russia exceeds the perceived distance of Jamaica to Cuba, plus that of Cuba to Russia - contrary to the triangle inequality." We can assume that perceived distance is either the same or approximately the same as dissimilarity, so that consistency with the triangle inequality requires.

$$\begin{aligned}&\text{Dissimilarity(Jamaica, Russia)}\\&\leq \text{Dissimilarity(Jamaica, Cuba)} + \text{Dissimilarity(Cuba, Russia)}\end{aligned} \tag{1}$$

Regarding the implications from this statement for similarity, we need a function that takes us from dissimilarity to similarity (or at least some indication of its properties) and Tversky does not provide this. Instead, he says "...the triangle inequality implies that if A is quite similar to B, and B is quite similar to C, then A and C cannot be very dissimilar from each other. Thus, it sets a lower limit to the similarity between A and C in terms of the similarities between A and B and between B and C." But, this expression is too vague to lead to a quantitative constraint. If one assumed that similarity is just the negative of dissimilarity, then one could write

$$\mathrm{Sim}(A, C) \geq \mathrm{Sim}(A, B) + \mathrm{Sim}(B, C) \tag{2}$$

but such an expression leaves us with some problems (e.g., we would need another function to take a negative, unbounded similarity measure to something that corresponds to e.g. similarity ratings; assuming the latter are closer to psychological similarity, in itself another assumption). No doubt some readers will find it unsatisfactory that a discussion, which is overall about similarity, actually is restricted to claims only about dissimilarity. But, for our purposes it is not necessary to resolve these issues, since we can easily formulate our discussion in terms of the inequalities based on dissimilarities above. With these points in mind, Tversky's example was as follows. Consider $A = Russia$ and $B = Jamaica$; $\mathrm{Dissimilarity}(Russia, Jamaica)$ is high. Consider also $C = Cuba$. But $\mathrm{Dissimilarity}(Russia, Cuba)$ is low (these countries are similar because of political affiliation) and $\mathrm{Dissimilarity}(Cuba, Jamaica)$ is also low (these countries are similar because of geographical proximity). Thus, Tversky's example suggests that

$$\begin{aligned} &\mathrm{Dissimilarity}(Russia, Jamaica) \\ &> \mathrm{Dissimilarity}(Russia, Cuba) + \mathrm{Dissimilarity}(Cuba, Jamaica) \end{aligned} \tag{3}$$

which suggests a violation of the triangle inequality. Interestingly, more elaborate theories of similarity, specifically developed to address Tversky's [1977] findings, do not always deal with violations of the triangle inequality straightforwardly (we will consider Krumhansl's, [1978] theory shortly, in Section 1.4.4).

1.3. *Diagnosticity*

The diagnosticity effect, a particular type of context effect, is another major finding from Tversky [1977]. Participants were asked to identify the country most similar to Austria, from a set of alternatives including Hungary, Poland, and Sweden. Participants typically selected Sweden. However, when the alternatives were Hungary, Sweden, and Norway, participants typically selected Hungary. Thus, the same similarity relation (e.g., the similarity between Sweden and Austria or the similarity between Hungary and Austria) appears to depend on which other stimuli are immediately relevant, showing that the process of establishing a similarity judgment may depend on the presence of other stimuli, not directly involved in the judgment. Tversky's [1977] explanation was that the diagnosticity effect arises from the grouping of some of the options. For example, when Hungary and Poland were both included, their high similarity made participants spontaneously code them with their obvious common feature (both were Communist bloc countries at the time), which, in turn, increased the similarity of the other two options, (Austria and Sweden) which were both Western democracies.

1.4. *Previous theoretical formalisms*

In the following sections we consider some significant previous theoretical efforts to account for Tversky's [1977] challenges. Such efforts have the same objective, but can vary widely in their assumptions, implementation, and structure, thus sometimes making it hard to identify their key distinguishing characteristics. Consideration of these previous theoretical approaches motivates our own proposal for a new approach, based on quantum theory.

1.4.1. *Extensions of the geometric model*

Let us first repeat the point that simple extensions of geometric models of similarity are unsatisfactory. In standard models [e.g. Shepard, 1980], the similarity between two entities A and B is given by

$$\mathrm{Sim}(A,B) = \exp(-c.\mathrm{distance}(A,B)) \tag{4}$$

where c is a constant. Clearly, such a function of similarity obeys symmetry. This basic definition could lead to an asymmetric similarity measure with the introduction of a directionality parameter, p_{AB}, indexed in a way to indicate that it may have a different value depending on whether we are considering the similarity of A to B or the similarity of B to A (see Nosofsky [1991], for these ideas). However, without a scheme for motivating particular values of the directionality parameter, this proposal cannot be said to explain asymmetry in similarity judgments a priori (even if it can post hoc reproduce the empirical results).

The basic geometric scheme also fails in the case of diagnosticity, since there is no mechanism by which to augment the computation of similarity for two entities by information for other, assumed relevant, objects. One could augment a basic similarity scheme with attentional weights, which could vary depending on contextual influences [cf. Nosofsky, 1984]. However, an approach like this would be incomplete without a precise understanding of how attentional weights can change, across different contexts. Overall, this simple extension of geometric models is a straw man and it is unsurprising that it fails. We will shortly see Krumhansl's [1978; see also Nosofsky, 1991] proposal.

1.4.2. *Extensions of the geometric model*

Tversky's [1977] Contrast Model proposed that

$$\mathrm{Sim}(A,B) = \theta f(A\cap B) - \alpha f(A-B) - \beta f(B-A) \tag{5}$$

where θ, α, β are constant parameters, $A \cap B$ denotes the common features between A and B, $A-B$ the features of A which B does not have and $B-A$ the features of B which A does not have [see also Bush & Mosteller, 1951; Eisler & Ekman, 1959]. Such a scheme can predict violations of symmetry if A has more features than B and the parameters α, β are different to each other and suitably set (e.g., $\theta > 0, \alpha = 1, \beta = 0$, allows the emergence of asymmetries from Tversky's contrast model, in the predicted direction). For example, regarding the Korea-China example, Tversky assumed that China has more features than Korea, because the average observer will know more about China than Korea. First we must assume that $\alpha > \beta$ in a directional judgment of similarity, so that distinctive features of the subject are more relevant than the distinctive features of the referent. Then, the similarity of Korea to China would be fairly high (a minimal negative contribution from $\alpha f(A-B)$, since Korea has very few features which China does not have). However, in comparing China to Korea, there is now a larger contribution from $\alpha f(A-B)$, which lowers the overall similarity result. Thus, according to Tversky's similarity model, China is predicted to be less similar to Korea than Korea is to China. Tversky's model of similarity is appealing, but still involves two independent parameters, which must have appropriate values to account for violations of symmetry. For example, if instead of assuming $\alpha > \beta$, we assume the reverse, then the model fails to predict the right direction for symmetry violation in the Korea-China example; There are some similarities between the α, β parameters in Tversky's similarity model and the directionality ones above [Nosofsky, 1991], in the sense that both kind of parameters are about defining a 'preferred' direction in similarity comparisons (that is, a direction that leads to higher similarities).

So, why did Tversky [1977] set the contrast model parameters one way, as opposed to another? Tversky's [1977] assumption was that when assessing $\mathrm{Sim}(A, B)$ then A is the subject and B is the referent and so "...the features of the subject are weighted more heavily than the features of the referent" [p. 333, Tversky, 1977]. This allows one to set $\alpha > \beta$, which enables possible violations of symmetry, as long as the compared objects differ in the number of distinctive features. While this assumption seems reasonable, it is also one which does not follow naturally from the rest of Tversky's [1977] model. Moreover, it is hard to logically exclude the alternative assumption (which leads to the exact opposite prediction), i.e., that it is the referent's features which are more heavily weighted. It is this assumption which basically allows the prediction of the asymmetry in a specific direction, so the extent to which it can be justified a priori goes hand in hand with our perception of whether the contrast model can explain the China, Korea asymmetry in an a priori way.

Tversky's [1977] contrast model provides the same elegant account for both the triangle inequality and the diagnosticity effect, in terms of how different contexts lead to the emergence of different diagnostic features (but see Krumhansl [1978], for some criticisms, relating to how weights are assigned to features, with varying contexts; her theory is considered in Section 1.4.4 below). His explanation for these empirical results is theoretically appealing, but some concerns can be expressed regarding the number and precise form of the emerging diagnostic features. In closing the discussion for Tversky's [1977] model, it is perhaps worth remarking that this detailed scrutiny of his work, so as to motivate the need for a new model (the quantum model), should not detract from the fact that his theory has had a profound and lasting influence on the similarity literature, probably more so than any other similarity theory.

1.4.3. *Classical probability theory*

In this section we consider whether classical (Bayesian) probability theory can provide an account of Tversky's [1977] challenges. It has to be said that classical probability theory is not obviously relevant to human similarity judgments. Nevertheless, cognitive models based on classical probability theory have been extremely successful in recent years [e.g., Griffiths et al., 2010; Oaksford & Chater, 2009; Tenenbaum et al., 2011] so it is worth exploring possible extensions in relation to similarity. The similarity between two instances could be modeled as the joint probability of both instances together,

$$Prob(A \wedge B) \tag{6}$$

Such a joint probability could be understood in terms of statements corresponding to both instances being concurrently true or in terms of the ease of having both thoughts together. We stress that our aim here is not to develop an operational model of similarity based on classical probability theory! Rather, we look at the Quantum Similarity Model and consider which operations are analogous in classical probability theory. The Quantum Similarity Model basically models similarities as conjunctive probabilities (the ease of having a thought about the first between two compared concepts and then the second). So, without worrying too much about operational details, we consider whether a similar approach might work with classical probability theory. However, the joint probability operator is symmetric in classical probability, so that

$$Prob(A \wedge B) = Prob(B \wedge A) \tag{7}$$

and so this scheme fails to account for violations of symmetry in similarity. Note that one could say that

$$Prob(A \wedge B | order1) \neq Prob(B \wedge A | order2) \tag{8}$$

but such a scheme offers a trivial solution to the problem of asymmetry (it is equivalent to the directionality parameter one above). Alternatively, one could model the similarity between two instances in terms of a conditional probability function, which can be asymmetric. In other words, one could postulate that

$$\mathrm{Sim}(A, B) = Prob(A|B) \neq Prob(B|A) = \mathrm{Sim}(B, A) \tag{9}$$

However, such a scheme does not work. Consider the paradigmatic Korea-China example again, from Tversky [1977], and assume that

$$\mathrm{Sim}(Korea, China) = Prob(China|Korea) \tag{10}$$

so that the similarity process involves assessing the probability of the second predicate given knowledge of the first (note, something like $Prob(China|Korea)$ could be interpreted as the conditional probability of thinking about China, given that we have been thinking about Korea). Then,

$$\begin{aligned} &\mathrm{Sim}(Korea, \neg China) = Prob(\neg China|Korea) \\ &= 1 - Prob(China|Korea) = 1 - \mathrm{Sim}(Korea|China) \end{aligned} \tag{11}$$

Since $\mathrm{Sim}(Korea|China)$ is assumed to be high, it follows that $\mathrm{Sim}(Korea|\neg China)$ has to be low. The latter conclusion seems reasonable, as all the predicates which satisfy $\neg China$ would, on average, have a low similarity to Korea (there are more countries which are dissimilar to Korea than ones which are similar). However, this approach can also lead to paradoxical predictions. Consider

$$\begin{aligned} &\mathrm{Sim}(Alaska, \neg China) = Prob(\neg China|Alaska) \\ &= 1 - Prob(China|Alaska) = 1 - \mathrm{Sim}(Alaska, China) \end{aligned} \tag{12}$$

Therefore, following this set of equalities in the reverse order, as $\mathrm{Sim}(Alaska, China)$ is very low, it must be the case that $\mathrm{Sim}(Alaska, \neg China)$ is very high. However, such a prediction seems counterintuitive.

Of course, classical probability theory is a sophisticated computational framework and it is possible that a satisfactory account of symmetry violations (and the rest of Tversky's [1977], challenges) can emerge. Our purpose here was to assess whether a basic classical probability model is consistent with violations of symmetry. This appears not to be the case. Moreover, it is not clear how this basic classical probability model could be extended in the case of the other relevant empirical results. A critic might note that (classical) probability theory has nothing to do with similarity judgments and this entire section is misguided. Nevertheless, the Quantum Similarity Model does exactly this: it provides a formalism in which probabilities (corresponding to the ease of having sequences of thoughts) lead to similarity judgments. Indeed, we think that approaching similarity judgments as probabilities (defined in a suitable way) is a worthwhile endeavor, insofar that this provides a framework for exploring commonalities between similarity and probabilistic inference [Shafir et al., 1990; Tversky & Kahneman, 1983].

1.4.4. *Krumhansl's distance-density model*

Krumhansl's [1978, 1988] distance-density model provides a principled extension to the basic geometric model of similarity. Her proposal rests on the assumption that alternatives lying within dense subregions of psychological space are subject to finer discrimination than alternatives lying in less dense subregions. With regards to similarity judgments, this implies that a pair of points a given distance apart in a dense region would have a lower similarity (greater psychological distance) as compared to an identical pair of points in a less dense region. More specifically, the distance between two points A and B in psychological space should be affected by the local density around each point, $D(A)$ and $D(B)$. The local density around a point reflects the number of other points within a certain radius. Thus,

$$d'(A,B) = d(A,B) + aD(A) + bD(B) \tag{13}$$

where $d(A,B)$ is the standard geometric distance, a and b are parameters that reflect the weight given to each density, and $d'(A,B)$ is the modified distance measure, as affected by local densities. As with Tversky's [1977] similarity model, it is immediately clear that if $a = b$ then

$$d'(A,B) = d(A,B) + a(D(A) + D(B)) = d'(B,A) \tag{14}$$

that is, unless $a \neq b$ no violations of symmetry are predicted. But, also as with Tversky's model, setting the parameters in different ways (e.g., $a, b > 0$, as in the original formulation of the model, versus $a, b < 0$) predicts asymmetries in different directions. In order to account for asymmetries, Krumhansl [1978] adopted an assumption equivalent to that of Tversky [1977], that is, that the density of one object influences the comparison more than the density of the other.

In the particular case of the Korea-China example, for a violation of symmetry to occur, one would need to assume that the local density around China is different from the local density around Korea. Krumhansl [1978] suggested that prominent objects are likely to have many features and so these objects are likely to share features with a greater number of other objects, as compared to objects with fewer features. Therefore, prominent objects are more likely to exist in denser regions of psychological space. Krumhansl's [1978] logic is perhaps intuitive, but it does raise some questions. For example, why should prominent objects (with more features) share a greater number of features with other objects? These additional features could be distinctive, as indeed is implied in Tversky's [1977] analysis.

Krumhansl's [1978] explanation for the triangle inequality is based on the idea that similarity judgments emphasize dimensions and features that objects have in common. As a result, stimuli which are far apart in the overall psychological space may be close to each other in a low dimensionality subspace, corresponding to the common dimensions between the stimuli. For example, Russia and Cuba are similar in the subspace of Communism, which corresponds to their common dimension. Krumhansl [1978, p.12] notes "Subspaces defined by obvious stimulus dimensions would seem to be likelier projections than subspaces not corresponding to such dimensions" and goes on to observe that such a scheme may be able to account for similarity relations inconsistent with the triangle inequality. However, this explanation involves some ad hoc assumptions. For example, why is similarity assessed in a subspace (such an assumption does not follow from her density model, nor is employed elsewhere; cf. Ashby & Perrin, [1988], why is the appropriate subspace determined in such a way etc. (these issues are similar to the corresponding criticism for Tversky's [1977] account).

The density model is easily consistent with the diagnosticity effect [Krumhansl, 1978, 1988]. Recall, the distance between two concepts increases as the density of one object increases. For example, $\mathrm{Sim}(Hungary, Austria)$ would change depending on the $D(Hungary)$ term, which is the density around Hungary. If we add Poland to the choice set, $D(Hungary)$ increases, the effective distance between Hungary and Austria also increases, and so the similarity between the two countries decreases. A perhaps unsatisfactory aspect of Krumhansl's [1978] account for the diagnosticity effect is that it fails to capture the (reasonable) intuition that different comparisons do evoke different relevant features (in Tversky's [1977], approach) or perspectives (in the quantum approach).

1.4.5. *Ashby & Perrin's general recognition theory*

Ashby and Perrin's [1988] general recognition theory is an established probabilistic approach to similarity for perceptual stimuli, also based on representations in a psychological space. In brief, the theory can readily account for violations of symmetry in similarity judgments of perceptual stimuli. Each time a stimulus is perceived it can correspond to a different point in psychological space, according to a particular probability distribution. Psychological space is divided into response regions, such that within each response region it is optimal to make a particular response. Thus, similarity between two stimuli depends on the extent to which the distribution of perceptual effects for the first stimulus overlaps with the optimal response region for the second stimulus. Formally, for a pair of two-dimensional stimuli A and B,

$$\mathrm{Sim}(A, B) = \iint_{R_B} f_A(x, y)dxdy \tag{15}$$

where R_A is the region in the x, y perceptual plane associated with response R_B and $f_A(x, y)$ is the probability density function for the distribution of perceptual effects of stimulus A (note that similarity is actually defined as a function of the above integral, but this is not relevant here). As Ashby and Perrin [1988] note, such a scheme can lead to violations of symmetry in a number of ways. For example, if stimulus B is associated with a greater response region than A then, in general, $\mathrm{Sim}(A, B) > \mathrm{Sim}(B, A)$ and if the perceptual effects distribution for A has a greater variability than X, then it is also the case that $\mathrm{Sim}(A, X) > \mathrm{Sim}(X, A)$.

As Ashby and Perrin [1988] observed, these intuitions can be related to the Korea-China example. First, because for many observers Korea will be a 'more vague and poorly defined concept' (p.133), the representation of Korea in psychological space will have a greater variability. Second, they argued that the response region for Korea would be smaller than that of China, because Korea is very similar to many other countries. According to general recognition theory, both these factors predict that $\mathrm{Sim}(Korea, China) > \mathrm{Sim}(China, Korea)$. But there are some problems with this account. Whether Korea or China is more similar to other countries is unclear. Ashby and Perrin [1988, p.133] note that "...for many people North Korea is very similar to several other countries." But, recall, Krumhansl [1978, p.454] made the exact opposite assumption, "If prominent countries...are those stimuli having relatively many features, then these objects have features in common with a larger number of different objects...". In other words, Krumhansl [1978] assumed that it is China, not Korea, which is similar to a greater number of other countries. Thus, Ashby and Perrin [1988] and Krumhansl [1978] make the exact opposite assumption, regarding whether it is Korea or China which is similar to a greater number of other countries. This shows the fickle nature of this assumption and how it can be (fairly easily) made one way or another, so that the corresponding models can describe an asymmetry in similarity judgments in the predicted direction and, equally easily, in the opposite direction as well.

Regarding the triangle inequality, Ashby and Perrin [1988] show how one can manipulate the perceptual effects distributions, so that two stimuli can be dissimilar to each other and yet both similar to a third stimulus, hence violating the triangle inequality. Such a situation can clearly be mapped to Tversky's [1977] Russia-Cuba-Jamaica example. One weakness of Ashby and Perrin's [1988] demonstration is that it appears to assume (see their Figure 4, p.133; one distribution circular on a plane, the other two elliptical, if one considers a suitable cross-section of the distributions) asymmetric and inequivalent perceptual effects distributions for the three stimuli. In the case of simple perceptual stimuli, presumably their form can be manipulated to produce arbitrary perceptual effects distributions. However, it is unclear whether such an assumption is reasonable in the case of, for example, comparisons between Russia, Cuba, and Jamaica. Why would the distributions for such countries have a different shape?

Regarding the diagnosticity effect, Ashby and Perrin's [1988] model can account for context effects on similarity judgments, in terms of how the presence of an additional stimulus C can modify the response region relevant in computing the similarity between two other stimuli, A and B. Specifically, the similarity between two stimuli A and C is

$$\text{Sim}(A, C) = \iint_{R_C} f_A(x, y) dx dy \tag{16}$$

Suppose that a third stimulus B is introduced near stimulus C. This means that the response region R_C is decreased, since a part of what used to be R_C is now the response region for R_B. Therefore, the integral

$$\text{Sim}(A, C) = \iint_{\text{New } R_C} f_A(x, y) dx dy \tag{17}$$

sums probability weight over a smaller area and so $\text{Sim}(A, C)$ is reduced. But, such a reduction of similarity between A and C is predicted regardless of where exactly a stimulus intermediate to A and C is introduced, as long as it is in between (in psychological space) A and C, and so leads to a reduction in the response region for C. In other words, with such a scheme, a 'diagnosticity' effect can emerge for stimuli without a corresponding natural grouping of some stimuli in psychological space, in contrast with the intuition in Tversky's [1977] empirical demonstration.

A more general issue with general recognition theory is that it is not a theory best suited for dealing with conceptual stimuli (a limitation which Ashby & Perrin, 1988, themselves acknowledged). For example, the argument for asymmetry in the Korea-China example or the diagnosticity effect also assumes that the decision boundary between response regions is optimal. Perhaps such an assumption is valid for perceptual stimuli studied across multiple repetitions but it is questionable as to whether it applies for one-shot similarity judgments between previously unencountered object pairs (Ashby & Perrin [1988], say that in one-shot cases additional assumptions can be made regarding the form of perceptual effects distribution). Moreover, the notion of confusability itself does not apply to most conceptual stimuli. Ashby and Perrin [1988] recognized this and provided a generalization to their similarity function, so that the overlap integral also includes a weighting term. Crucially, their generalized similarity function still implies that similarity depends on the proximity of the perceptual effects distributions, and this proximity is likely to be low for many conceptual object pairs (since it is rarely the case that we can confuse one real-world object for another, even for objects which are quite similar, such as apples and pears). Thus, the general recognition theory guides us to a prediction of universally low similarity in the case of pairs of conceptual objects. Overall, the key strength of general recognition theory is that a researcher can produce predictions regarding the classification of simple, perceptual stimuli, on the basis of precise manipulations of the perceptual effects distribution for each stimulus. In such cases, the general recognition theory is probably the best of the available theories. However, applying this approach to the case of conceptual stimuli, such as the ones in Tversky's [1977] challenges, leads to difficulties.

We complete our short review by a (fairly obvious, we hope) qualification: our review was extremely selective, focusing primarily on the formal models, which have emerged as major candidates for explaining the key findings from Tversky [1977]. It is important to bear in mind that there have been other, influential theoretical perspectives for these results, not based on theory specified in mathematical terms [especially in relation to asymmetries, e.g., Bowdle & Gentner, 1997; Bowdle & Medin, 2001; see also Gleitman et al., 1996]. Moreover, there has been extensive work on various relevant methodological aspects of how asymmetries in similarity and the other relevant effects are demonstrated [e.g., Aguilar & Medin, 1999]. Note, however, most researchers currently do accept the reality of most of these effects.

2. The Quantum Similarity Model (QSM)

Next, we present an alternative model for similarity judgments based on Quantum Probability (QP) theory [Pothos et al., 2013; Pothos & Trueblood, 2015]. QP theory is a theory for how to assign probabilities to events [Hughes 1989; Isham 1989], alternative from classical probability theory. We call QP the rules for how to assign probabilities to events from quantum mechanics, without any of the physics. QP has the potential to be relevant in any area of science, where there is a need to formalize uncertainty. Regarding psychology, clearly, a major aspect of cognitive function is the encoding of uncertainty and therefore QP is potentially applicable in cognitive modeling. QP theory and classical probability theory are founded from different sets of axioms and so are subject to alternative constraints. The use of QP for modeling cognitive processes follows on from a number of recent attempts to describe various phenomena in psychology, and the social sciences more generally, using non-classical models of probability. Certain types of cognitive processing, in situations where it appears there may be incompatibility between the available options [Busemeyer et al, 2011], may be better modeled using QP theory.

The QSM follows the recent interest in the application of QP theory to cognitive modeling. Applications of QP theory have been presented in decision making [White et al., 2014; Busemeyer, Wang, & Townsend, 2006; Busemeyer et al., 2011; Bordley, 1998; Lambert, Mogiliansky, Zamir, & Zwirn, 2009; Pothos & Busemeyer, 2009; Trueblood & Busemeyer, 2011; Yukalov & Sornette, 2010]; conceptual combination [Aerts, 2009; Aerts & Gabora, 2005; Blutner, 2008]; memory [Bruza, 2010; Bruza et al., 2009], and perception [Atmanspacher, Filk, & Romer, 2004]. For a detailed study on the potential of using quantum modeling in cognition see Busemeyer and Bruza [2011] and Pothos and Busemeyer [2013].

A unique feature of the QSM is that, whereas previous models would equate objects with individual points or distributions of points, in the quantum model objects are entire subspaces of potentially very high dimensionality. This is an important generalization of geometric models of similarity, as it leads to a naturally asymmetric similarity measure. We first present an outline of the QSM and its main features. Subsequently, we consider again the violations of symmetry, triangle inequalities and the diagnosticity effect, from Tversky [1977], and how the QSM helps provide relevant explanations.

2.1. *A new psychological space*

Representations in QP theory are based on a multidimensional space. These representations are geometric ones, but such that the represented entities (stimuli, concepts, etc.) are not just single points in a geometric space, but rather entire subspaces. This provides a very natural approach to the problem of capturing differences in knowledge: the more you know about something (stimuli, concepts, etc.) the greater the dimensionality of the subspace for that entity. Thus, QT theory provides a unique, novel way to approach representation, that extends previous efforts both in psychology [Shepard, 1987] and generally [cf. Kintsch, 2014]. Furthermore, the idea of an overlap between vectors and subspaces as a measure of similarity has a long history in psychology [Sloman, 1993]; QP theory provides a more principled approach to this idea.

The QSM is based on a Hilbert space H, which is a complex vector space (with some additional properties), that represents the space of possible thoughts. The overall space can be divided into (vector) subspaces representing particular concepts. Imagine a concept A. The subspace corresponding to this concept is associated with a projection operator P_A. Note that, in general, suitable spaces for modeling similarity judgments would be of very high dimensionality. However, in specific experimental situations, low dimensionality spaces usually provide adequate approximations.

In quantum models, in general, the current state of the system is given by a density operator ρ on H or, where simplifying conditions apply, a state vector. In psychological applications, including in the QSM, the state of the system corresponds to whatever a person is thinking at a particular time. More specifically, in the QSM, the relevant state is the mental state of a participant, just prior to a similarity judgment. Note, the state vector will often be at an angle to the various subspaces in the Hilbert space and it is determined by, for example, the experimental instructions; in other cases, the state vector may represent the expected degree of knowledge of participants. By projecting this current state onto the different subspaces of the relevant Hilbert space and then computing the squared length of the projected vector, we have a measure of the consistency between the state vector and the other entities represented in our quantum space. Below, (Figure 1) we will see a graphical illustration for how to compute these operations.

The QSM is a departure from classical geometric representation schemes. It offers a rigorous framework for associating concepts with subspaces and it provides us with representational flexibility, in that there are no constraints in the number of features one can employ for representing different concepts (within the same application, once can have subspaces varying greatly in dimensionality). Note, in a classical representational approach based on psychological spaces, each object must be represented with the same number of dimensions (all the available ones).

Consider next how to compute the similarity between two concepts in the QSM. The similarity between two concepts A and B is computed as

$$\mathrm{Sim}(A, B) = \mathrm{Tr}(P_B P_A \rho P_A) \tag{18}$$

where ρ is the mixed knowledge state of the system; if the knowledge state is pure, $\rho = |\psi\rangle\langle\psi|$, the expression for similarity reduces to

$$\mathrm{Sim}(A, B) = |P_B P_A |\psi\rangle|^2 \tag{19}$$

One of the important parts of the model is how to specify the current state of system or the knowledge state vector, as we called it before. We discuss this shortly.

We are going to follow the China-Korea example from Tversky [1977] to explain how the subspaces of the Hilbert space should be specified in our model. China would correspond to a subspace of the relevant knowledge space and Korea would correspond to another subspace. A subspace could be a ray spanned by a single vector, or a plane spanned by a pair of vectors, or a three dimensional space spanned by three vectors, etc. In this example, we represent China as a subspace spanned by two orthonormal vectors, $|v_1\rangle$ and $|v_2\rangle$, that is, the China subspace is two-dimensional and $|v_1\rangle$ and $|v_2\rangle$ are *basis* vectors for the China subspace. All the vectors of the form $a|v_1\rangle + b|v_2\rangle$ where $|a|^2 + |b|^2 = 1$ (as is required for a state vector in quantum theory) represent the concept of China. The concept of China itself is about lots of things. For example, when we think about China we think about culture, food, language, etc. To represent China as a subspace means that all these thoughts and properties, of the form $a|v_1\rangle + b|v_2\rangle$, are consistent with this concept and are contained within the China subspace. Here, we can see a key feature of the QSM, and at the same time, some commonalities with other models of psychological similarity, that is, that concepts correspond to regions of psychological spaces [Ashby & Perrin, 1988; Grdenfors, 2000; Nosofsky, 1984]. Further, imagine that we want to represent the idea that we have a greater knowledge for China than for Korea. We would represent China as a two dimensional space (we have a greater range of thoughts/properties/statements) and Korea as a one dimensional space.

Let us note that a thought of the form

$$|\psi\rangle = a|v_1\rangle + b|v_2\rangle \tag{20}$$

is neither about $|v_1\rangle$ nor $|v_2\rangle$, but rather reflects the potentiality that the person will end up definitely thinking about $|v_1\rangle$ or $|v_2\rangle$. In QP theory we cannot assign definite meaning to superposition states such as $a\left|v_1\right\rangle+b\left|v_2\right\rangle$. This is a result of the Kochen-Specker theorem. If $|a| > |b|$, this means that the person has a greater potential to think of $\left|v_1\right\rangle$ than $\left|v_2\right\rangle$. The mathematical expression for the concept of China would be a projector denoted as

$$P_{China} = \left|v_1\right\rangle\left\langle v_1\right| + \left|v_2\right\rangle\left\langle v_2\right| \tag{21}$$

Therefore, the mathematical expression of the collection of thoughts about China $|\psi\rangle$ is that

$$P_{China}\left|Thought\right\rangle = \left|Thought\right\rangle \tag{22}$$

so, the collection of these vectors represents, in the QSM, the range of thoughts consistent or part of the concept. For example, if we think about Chinese food, then

$$|\psi\rangle = \left|Chinese\,Food\right\rangle \tag{23}$$

and

$$P_{China}|\psi\rangle = P_{China}\left|Chinese\,Food\right\rangle = \left|Chinese\,Food\right\rangle \tag{24}$$

showing that this is a thought included in the China concept. How are we to determine the set of appropriate vectors, properties, or dimensions, especially given that different subsets of properties of a particular concept are likely to correlate with each other? This is an issue common to all geometric approaches to similarity. Recent work, especially by Storms and collaborators [e.g., De Deyne et al., 2008], shows that this challenge can be overcome, for example, through the collection of similarity information across several concepts or feature elicitation. Then, the relatedness of the properties will determine the overall dimensionality of the concept.

In the next section we discuss how to compute similarity in our QSM.

2.2. *Computing similarity*

In QP theory, to examine the degree to which the state vector is consistent with the subspace we need to employ a projector. We need (1) a particular subspace, which is China in our case and (2) a suitable knowledge state vector (or, more generally, a density matrix). A projector can be represented by a matrix, which takes a vector and projects it (lays it down) onto a particular subspace. In other words, (2) has to be projected into (1). Let us illustrate this in Figure 1, where we can see how we project vector B onto vector A; note, both vectors are unit length. We represent in red the projection, which would be another vector that corresponds to the part of B which is contained in A.

Mathematically, this is denoted by $|A\rangle\langle A|B\rangle$, noting that $P_A = |A\rangle\langle A|$ is the projector onto the A ray. Indeed, the notation $|A\rangle\langle A|B\rangle$ indicates a multiplication between a vector and an inner product. But, from elementary geometry we have that the inner product between two real vectors is $\langle A|B\rangle = |A|.|B|.\cos(\theta)$, where θ is the angle between the two vectors (see also Sloman, 1933]. If the two vectors are normalized, then $\langle A|B\rangle = \cos(\theta)$.

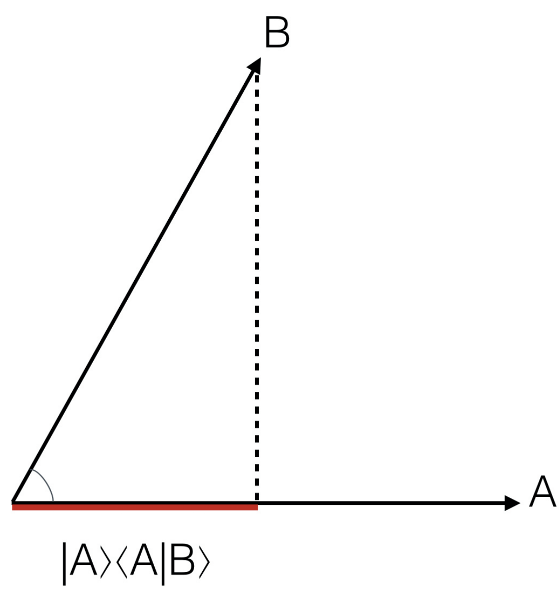

Fig. 1. Illustration of the projection of vector B onto vector A.

Let us follow the same procedure following the China example above. The projector onto the China subspace is denoted by P_{China}. Then, to compute the part of the vector $|\psi\rangle$ that is contained in the China subspace we need to compute the projection $P_{China}|\psi\rangle$. To compute the probability that the state vector is consistent with the corresponding subspace, we need to compute the length of the projection squared. The probability that a thought is consistent with the China concept equals

$$|P_{China}\,|\psi\rangle\,|^2 = \langle\psi|\,P_{China}\,|\psi\rangle \tag{25}$$

If the state vector is orthogonal to a subspace, then the probability is 0. This can also be written as

$$p(China) = \langle China\rangle_{\rho} = \mathrm{Tr}(P_{China}\rho) \tag{26}$$

if the initial state is a density matrix ρ, instead of a pure state $|\psi\rangle$. Thus, the probability that the initial knowledge state is consistent with the concept China is given as a measure of the overlap between the knowledge state and the subspace.

The QSM proposes that the similarity between two concepts is determined by the sequential projection from the subspace corresponding to the first concept to the one for the second concept. In other words, making a similarity judgment or comparison is a process of thinking about the first of the compared concepts, followed by the second. The similarity between Korea and China may therefore be written as,

$$\mathrm{Sim}(Korea, China) = \mathrm{Tr}(P_{China}P_{Korea}\rho P_{Korea}) \tag{27}$$

or

$$\mathrm{Sim}(Korea, China) = |P_{China}P_{Korea}\,|\psi\rangle\,|^2 \tag{28}$$

depending on whether the initial state is a density matrix ρ or a pure state $|\psi\rangle$.

2.3. *Reproducing asymmetries*

As just noted, in the QSM, similarity between two concepts A and B is defined as $\mathrm{Sim}(A, B) = |P_B P_A\,|\psi\rangle\,|^2$, that is, a process of thinking about concept A first and concept B second. Critically, the term $|P_B P_A\,|\psi\rangle\,|^2$ depends on four factors. First, how the initial state is set. In the case of comparing two concepts, we think the most plausible assumption is that ψ is set so that it is neutral/unbiased, between A and B. Second, similarity judgments are often formulated in a directional way [Tversky, 1977]. When this is the case, we suggest that the directionality of the similarity judgment determines the directionality of the sequential projection, i.e., the syntax of the similarity judgment matches the syntax of the quantum computation. Thus, there is a mechanism which potentially allows asymmetries in similarity judgments, when the projectors corresponding to the compared concepts do not commute (this will be the case, in general). Third, of course it depends on the angle between the subspaces. Finally, it depends on the relative dimensionality of the subspaces for concepts A and B (recall, greater dimensionality means greater knowledge).

In this section we are interested in how asymmetries can emerge from the QSM. We compare $|P_{Korea}P_{China}\,|\psi\rangle\,|^2$ and $|P_{China}P_{Korea}\,|\psi\rangle\,|^2$, noting that in both cases, the state vector is set so that it is neutral between the concepts compared, China and Korea. Note that

$$|P_{Korea}P_{China}\,|\psi\rangle\,|^2 = |P_{Korea}\,|\psi_{China}\rangle\,|^2 |P_{China}\,|\psi\rangle\,|^2 \tag{29}$$

and

$$|P_{China}P_{Korea}\,|\psi\rangle\,|^2 = |P_{China}\,|\psi_{Korea}\rangle\,|^2 |P_{Korea}\,|\psi\rangle\,|^2 \tag{30}$$

where $|\psi_{China}\rangle = P_{China}\,|\psi\rangle\,/|P_{China}\,|\psi\rangle\,|$ and $|\psi_{Korea}\rangle$ are normalized (length=1) vectors in the corresponding subspaces. Note also that, by assumption, $|P_{Korea}\,|\psi\rangle\,|^2 = |P_{China}\,|\psi\rangle\,|^2$, which is the condition that the mental state vector is unbiased between the two concepts. Then, the similarity between Korea and China vs. China and Korea reduces to comparing

$$|P_{China}\,|\psi_{Korea}\rangle\,|^2 \quad \text{(similarity of Korea to China)} \tag{31}$$

and

$$|P_{Korea}\,|\psi_{China}\rangle\,|^2 \quad \text{(similarity of China to Korea)} \tag{32}$$

But, in the former case, we project a vector to a higher dimensionality subspace, than in the latter. Thus, in the former case, there is more opportunity, so to say, to preserve the vector's amplitude. Thus, in the former case, the projection will (on average) have greater length.

In Figure 2, we illustrate the relevant subspaces and projections. The green line corresponds to a one-dimensional subspace (Korea), the blue plane to a two-dimensional subspace (China), and the black line to the state vector (set in such a way that it is neutral between the two subspaces, as postulated by the QSM). The length of the first projection corresponds to a solid red line and, by assumption, is the same regardless of whether we project to the ray or onto the plane. But, the length of the second projection, illustrated as a yellow line, differs depending on whether it is to a ray or to a plane, so that when this second projection is onto the plane, it is longer. Panel (a) shows a process of thinking about Korea first and then China, that is, $P_{China}P_{Korea}$;

$$\text{Sim}(Korea, China) = |P_{China}P_{Korea}\,|\psi\rangle\,|^2 \tag{33}$$

which is the squared length of the yellow line. Analogously for panel (b). This illustrates how

$$|P_{China}P_{Korea}\,|\psi\rangle\,|^2 > |P_{Korea}P_{China}\,|\psi\rangle\,|^2 \tag{34}$$

that is, the square of the yellow line in panel (a) is greater than the square of the yellow line in panel (b).

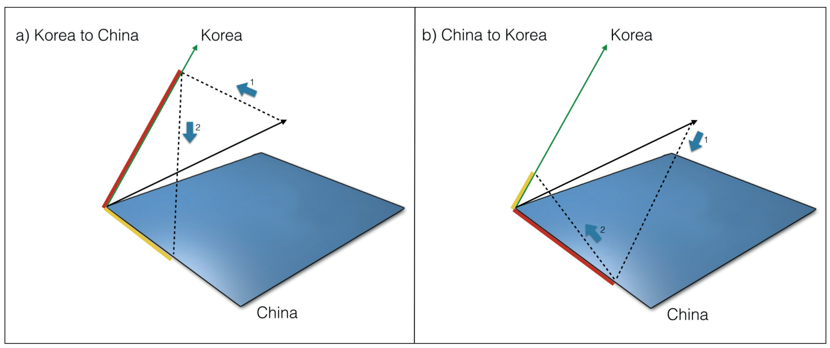

Fig. 2. Illustration for how to compute Sim(Korea, China) and Sim(China, Korea) using the QSM.

As extensively discussed in Pothos et al. [2013], the QSM thus allows a prediction of asymmetry in the case of the Korea, China example (and, obviously, all cases where there is a difference in degree of knowledge) to emerge naturally. As we noted above, in order to generate these asymmetries we need some principle for fixing the initial state. Usually we will (partly) fix the initial knowledge state by demanding that it is unbiased, that is, that there is equal prior probability that the initial state is consistent with either, say, Korea or China. Such an assumption is analogous to that of a uniform prior in a Bayesian model. Then, it is straightforward to show that

$$\mathrm{Sim}(Korea, China) \sim |P_{China}\,|\psi_{Korea}\rangle\,|^2 \tag{35}$$

whereby the vector $|\psi_{Korea}\rangle$ is a normalized vector contained in the Korea subspace. Therefore, the quantity $|P_{China}\,|\psi_{Korea}\rangle\,|^2$ depends on only two factors, the geometric relation between the China and the Korea subspaces and the relative dimensionality of the subspaces.

2.4. *Reproducing violations of the triangle inequality*

The QSM leads to violations of the triangle inequality in a way similar to how Tversky [1977] suggested such effects arise. As our representations are subspaces, different regions in the overall space end up reflecting the features characteristic of the corresponding concepts. We will follow another of Tversky's experiments as an example (Figure 3). We have a Hilbert space with Russia (in blue), Cuba (in red) and Jamaica (in green). All of Russia, Cuba, and Jamaica are represented as one dimensional subspaces, for simplicity. The region between Russia and Cuba will overall reflect the property of communism, noting that both countries are consistent with this property. Next, we can imagine a different region to the communist one containing Cuba and Jamaica. The shared characteristic of Cuba and Jamaica is their geographical proximity (they are both in the Caribbean), so this second region will likewise correspond to this property. It should be hopefully straightforward to then see how, if Cuba is on the boundary of the communism and Caribbean regions in psychological space, we can have Cuba highly similar to Russia (represented as (1) in dashed lines in Figure 3), Cuba highly similar to Jamaica (represented as (2)), but Russia and Jamaica dissimilar from each other (represented as (3)), thus violating the triangle inequality, i.e., producing

$$\begin{aligned}\text{Dissimilarity}(Russia, Cuba) + &\text{Dissimilarity}(Cuba, Jamaica) \\ &< \text{Dissimilarity}(Russia, Jamaica)\end{aligned} \tag{36}$$

2.5. *Reproducing diagnosticity*

The diagnosticity effect is central in the debate on whether distance-based similarity models are adequate or not and in this section we will show how the QSM can accommodate context when computing similarity judgments.

Sometimes what we think just prior to a comparison may be relevant to the comparison itself. Therefore, when computing the similarity of A and B we have to take into account the influence of some contextual information, C. As in all other computational examples we have seen, in the QSM C has to be represented by a subspace. Following Tversky's [1977] diagnosticity effect experiment, this information C could correspond to the alternatives in the task he employed. The similarity between A and B should then be computed as,

$$\text{Sim}(A, B) = |P_B P_A \left|\psi'\right\rangle|^2 = |P_B \left|\psi'_A\right\rangle|^2 |P_A \left|\psi'\right\rangle|^2 \tag{37}$$

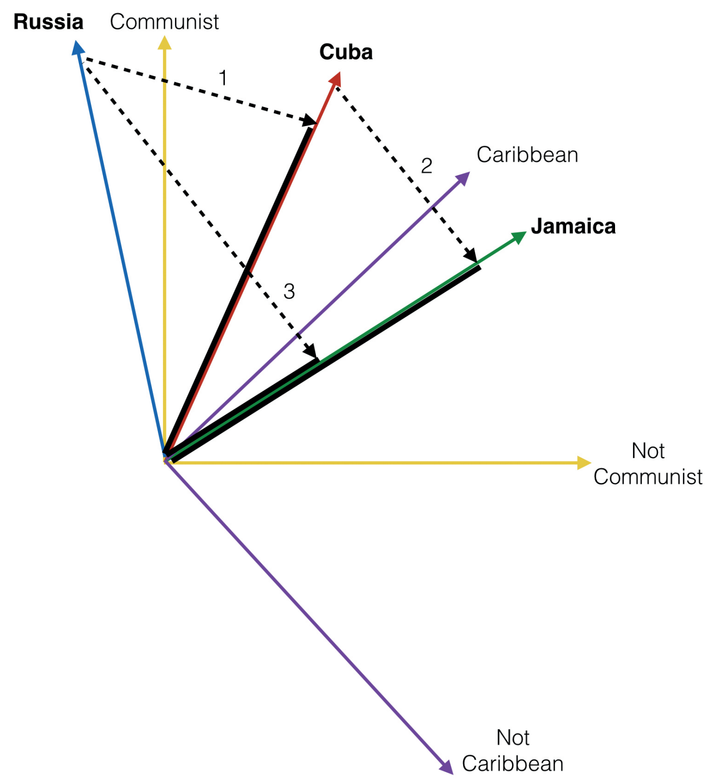

Fig. 3. An illustration of how the QSM can accommodate Tversky's [1977] finding, which is often interpreted as a violation of the triangle inequality.

where $|\psi'\rangle = |\psi_C\rangle = P_C |\psi\rangle / |P_C |\psi\rangle|$ is no longer a state vector neutral between A and B, but rather one which reflects the influence of information C. If we minimally assume that the nature of this contextual influence is to think of C, prior to comparing A and B, then

$$\begin{aligned} \mathrm{Sim}(A, B) = |P_B P_A |\psi'\rangle|^2 &= |P_B P_A (P_C |\psi\rangle)/|P_C |\psi\rangle||^2 \\ &= |P_B P_A P_C |\psi\rangle|^2 / |P_C |\psi\rangle|^2 \end{aligned} \tag{38}$$

In other words, if we first think about A and then about B when making a similarity comparison between A and B, then in the context of some other information C should involve an additional first step of first thinking about

C. Computationally, we prefer to employ $|P_B P_A P_C |\psi\rangle|^2$, since

$$|P_B P_A P_C |\psi\rangle|^2 = |P_B P_A |\psi_C\rangle|^2 |P_C |\psi\rangle|^2 = |P_B |\psi_{AC}\rangle|^2 |P_A |\psi_C\rangle|^2 |P_C |\psi\rangle|^2 \tag{39}$$

where $|\psi_C\rangle = P_C |\psi\rangle / |P_C |\psi\rangle|$ and $|\psi_{AC}\rangle = P_A |\psi_C\rangle / |P_A |\psi_C\rangle|$. Therefore, the similarity comparison between A and B is now computed in relation to a vector which is no longer neutral, but contained within the C subspace. Depending on the relation between subspace C and subspaces A and B, contextual information can have a profound impact on a similarity judgment.

As we have done in the previous sections, we present an illustration (Figure 4) of how the diagnosticity effect arises from the QSM, using Tversky's [1977] example. As we explained in a previous section, in his experiment participants had to identify the country most similar to a particular target, from a set of alternatives, and the empirical results showed that pairwise comparisons were influenced by the available alternatives. Specifically, participants were asked to decide which country was most similar to Austria, amongst a set of candidate choices. When the alternatives were Sweden, Poland, and Hungary, most participants selected Sweden (49%), so implying that Sim($Sweden, Austria$) was the highest (panel (a) in Figure 4) . When the alternatives where Sweden, Norway, and Hungary, Hungary was selected most frequently (60%), not Sweden (14%). Thus, changing the range of available alternatives can apparently radically change the similarity between the same two alternatives. Tversky's [1977] explanation for this result was that the range of alternatives led to the emergence of different diagnostic features (either 'Eastern European' countries or 'Scandinavian' countries), which in turn impacted on the similarity judgment. Analogous demonstrations were provided with schematic stimuli. Figure 4 shows a plausible QSM arrangement for Austria, Sweden, Poland, Norway and Hungary and the corresponding projections that lead to the diagnosticity effect.

The QSM is able to reproduce the main empirical findings from Tversky's [1977] diagnosticity effect experiment and this approach also leads to qualitative predictions about when the effect is likely to be present or absent, based on the geometric relationships between the stimuli in psychological space. Nevertheless, regarding the emergence of the diagnosticity effect, the QSM involves a number of assumptions worth evaluating in detail. These assumptions concern mainly the way the context items influence the similarity judgment and the role of the initial knowledge state.

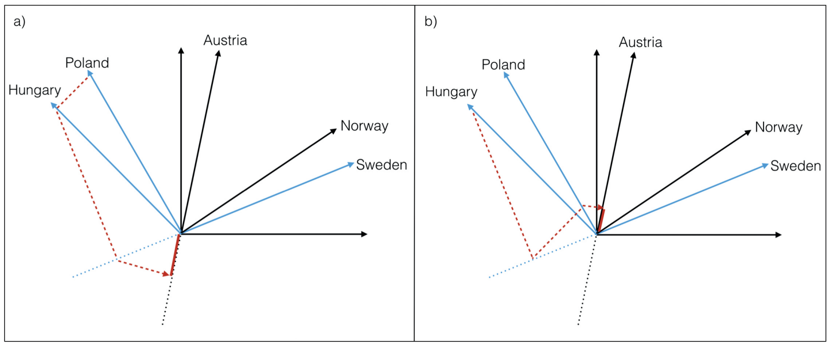

Fig. 4. An illustration of how QSM can account for the results from Tversky's [1977] diagnosticity experiment. The order of projection on panel (a) is such that we start from the context elements, Poland (arbitrarily chosen first in the illustration), then Hungary, then Sweden, then Austria. If we were computing a similarity without context, we would just have a projection from Sweden to Austria, to correspond to the similarity of Sweden (first projection) to Austria. Analogously for panel (b).

Let us first consider how the diagnosticity effect emerges in the QSM. As discussed above, context corresponds to successive projections between the context elements. When the context elements are grouped together (as for Hungary, Poland), projecting across them leads to little loss of amplitude of the state vector, so that the similarity judgment ends up being higher. When there is no grouping across any of the possible contexts, then the effect of context is simply to uniformly scale the similarity judgments. Thus, context can make the same similarity comparison appear higher or lower, depending exactly on the grouping of the context elements [Pothos et al., 2013]. The intuition for how the quantum model produces the diagnosticity effect is thus not much different from that of Tversky's [1977]. But, in Tversky's [1977] model, it has to be assumed that diagnostic features are 'invoked', as a result of grouping, while in the quantum model, the diagnosticity effect emerges directly from the presence of a grouping. In the next section, we address some challenges regarding how the QSM reproduces diagnosticity effects, the kind of novel predictions that the model can produce and alternative motivations for some of the QSM assumptions.

3. Conclusions, Challenges and Further Directions

The objective of this chapter was to present the QSM and consider how it can account for Tversky's [1977] key challenges. The QSM generalizes the notion of geometric representations, but the emergent similarity metric is not distance-based, thus avoiding many of the criticisms Tversky [1977] made against distance-based similarity models. The QSM can be seen as an example of a new way of thinking about cognitive modeling, that may also be applied to constructive judgments [i.e. White et al., 2014], belief updating and many other analogous areas of research.

The QSM was developed to associate knowledge with subspaces. This idea of representations as subspaces allows us to capture the intuition that a concept is the span of all the thoughts produced by combinations of the basic features that form the basis for the concept. The QSM also helped us to cover some key empirical results: basic violations of symmetry, violations of the triangle inequality and the diagnosticity effect, all from Tversky [1977]. Nevertheless, we offer below a list of challenges for the QSM and open issues for further research (some of which we are in the process of addressing). It is important to establish whether the QSM model makes any novel predictions about similarity judgments in particular cases. These could either take the form of new qualitative effects or of quantitatively accurate predictions for similarity judgments. Our overall conclusion is that further work is clearly needed with the QSM, though the new results are encouraging for the overall potential of the approach.

3.1. *Fixing the initial state*

One problem with the QSM, as presented, is that it relies on a particular choice of initial state in order to reproduce the asymmetry/diagnosticity effects. Even in set-ups where one can partially fix the initial state by demanding it to be unbiased, this typically leaves some degrees of freedom unfixed (that is, this requirement does not always produce a unique state vector; there are equivalent neutral state vectors and it is unclear why one would prefer one option, as opposed to another). Further research is needed in terms of determining in a reliable way how to set the knowledge state vector for a participant or a group of participants. Moreover, we noted that the state vector could be affected by information relevant to the similarity judgment. In the diagnosticity effect example, there is a specific procedure for incorporating relevant effects, but we would like a more general scheme for how relevant prior information impacts on the state vector.

3.2. *Interpreting the subspaces*

More work is needed concerning the interpretation of the dimensions of the subspaces, which represent each concept (or stimulus, etc.). The dimensions of each subspace have to correspond to the independent feature/characteristics, which collectively capture our knowledge of a concept. As an example, consider the standard Cartesian xyz coordinate system. There are many vectors which are in between the xyz coordinates. However, we can represent all this information, in terms of coordinates just along the three main axes (xyz) of the overall space. Likewise, when considering the subspace representing e.g. China, there are going to be many characteristics which highly correlate with each other. For example, our knowledge of Chinese art and culture relates to our knowledge of Chinese language etc. So, for a particular concept, we may have a greater or smaller number of individual features, but the extent to which the dimensionality of the corresponding subspace will be greater or smaller depends on the relatedness of the features. Regarding the emergence of asymmetries in similarity judgments, this is the main difference between Tversky's [1977] thinking and the QSM: the former predicts asymmetries in terms of the number of features, the latter as some function of the number of independent features. It is clearly desirable to empirically examine this difference in prediction.

3.3. *Modeling context effects*

One important challenge in developing the QSM is further formalizing the way in which contextual influences are taken into account. The idea of incorporating context as prior projections works well, but can the QSM be extended such that these prior projections can be motivated in a more rigorous way?

A great focus for further work with the QSM concerns the diagnosticity effect. This effect has proved difficult to replicate [e.g., see Evers & Lakens, 2014] and it would be interesting to see whether the QSM could generate any new predictions, regarding the emergence or suppression of the diagnosticity effect. We are interested in exploring whether the QSM model can provide insight into why the diagnosticity effect has proved elusive in its replicability. The diagnosticity effect is also significant because the quantum formalism, overall, is often said to embody strong contextual influences. So, perhaps, quantum theory would be particularly suitable for modeling context effects in similarity judgments? The diagnosticity effect does emerge fairly naturally from the QSM, but the mechanisms that allow this are not the traditional contextual mechanisms in quantum theory (e.g., relating to entanglement or incompatibility). The difficulty lies in the fact that contextual influences in similarity appear to arise depending on the degree of grouping of some of the options in the relevant choice set. The QSM is sensitive to the grouping of the context elements, but there is still a challenge to embed the contextual mechanism in the QSM within a more rigorous, formal framework.

3.4. *Dealing with frequency and prototypicality*

An important gap in the QSM concerns how to deal with asymmetries arising from differences in the frequency of presentation of stimuli [Polk et al., 2002] or from differences in prototypicality [Rosch, 1975]. This failure is interesting when we note that there appears to be an obvious way to include such effects. Presumably what distinguishes a prototypical stimulus from a non-prototypical one, or a stimulus presented many times from one presented only infrequently, is the increased potentiality for a participant to think about this stimulus. It would be interesting to see how the QSM could account for how differences in frequency/prototypicality can lead to asymmetries in similarity.

3.5. *Analogical similarity judgments*

Another important focus concerns so-called analogical similarity judgments [e.g., Gentner. 1983; Goldstone, 1994; Larkey & Love, 2003]. Analogical similarity is a vast topic and here we focus on one aspect of it, namely the idea that, for example, when comparing two people, Jim and Jack, if they both have black hair, this will increase their similarity, but if Jim has black hair and Jack has black shoes (and blond hair), this will have less impact on their similarity. That is, work on analogical similarity recognizes that objects often consist of separate components. Commonalities on matching components (e.g., black hair) increase similarity more so than commonalities on mismatching components (e.g., black hair and black shoes). It is currently unclear whether there is a genuine distinction between cognitive processing corresponding to basic similarity tasks [as in Tversky, 1977] and analogical similarity ones (some researchers have suggested that different cognitive systems may mediate the two types of judgments; Casale et al. [2012]). Nevertheless, there have been largely separate corresponding literatures for these two kinds of similarity judgments, with different objectives.

We think that the QSM can be extended to incorporate analogical similarity, because quantum theory already has extensive machinery in place for combining individual components into a whole [cf. Smolensky, 1990]. Indeed, we have been pursuing an approach based on tensor products [Pothos & Trueblood, 2015].

Acknowledgements

EMP and JMY were supported by Leverhulme Trust grant RPG-2013-00. Further, EMP was supported by Air Force Office of Scientific Research (AFOSR), Air Force Material Command, USAF, grants FA 8655-13-1-3044. JST was supported by NSF SES-1326275. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

References

1. Aerts, D., & Gabora, L. (2005). A theory of concepts and their combinations II: A Hilbert space representation. Kybernetes, 34, 192-221.
2. Aerts, D. (2009). Quantum structure in cognition. Journal of Mathematical Psychology, 53, 314-348.
3. Aguilar, C. M., & Medin, D. L. (1999). Asymmetries of comparison. Psychonomic Bulletin & Review, 6, 328-337.
4. Ashby, G. F. & Perrin, N. A. (1988). Towards a Unified Theory of Similarity and Recognition. Psychological Review, 95, 124-150.
5. Atmanspacher, H., Filk, T., & Romer, H. (2004). Quantum Zeno features of bistable perception. Biological Cybernetics, 90, 33-40.
6. Blutner, R. K. (2008). Concepts and Bounded Rationality: An Application of Niestegge's Approach to Conditional Quantum Probabilities. In L. Accardi, et al. (eds.), Foundations of Probability and Physics 5, Vol. 1101, pp. 302-10. American Institute of Physics Conference Proceedings. New-York.
7. Bordley, R. F. (1998). Quantum mechanical and human violations of compound probability principles: Toward a generalized Heisenberg uncertainty principle. Operations Research, 46, 923-926.
8. Bowdle, B. F. & Gentner, D. (1997). Informativity and asymmetry in comparisons. Cognitive Psychology, 34, 244-286.
9. Bowdle, B. F. & Medin, D. L. (2001). Reference-point reasoning and comparison asymmetries. In J. D. Moore & K. Stenning (Eds.) Proceedings of the 23rd Annual Conference of the Cognitive Science Society, pp. 116-121. Psychology Press.
10. Bruza, P. D. (2010). Quantum Memory. Australasian Science, 31, 34-35.
11. Bruza, P. D., Kitto, K., Nelson, D. & McEvoy, C. L. (2009). Is there some-

thing quantum-like about the human mental lexicon? Journal of Mathematical Psychology, vol. 53, pp. 362-377.

12. Busemeyer, J. R. & Bruza, P. (2011). Quantum models of cognition and decision making. Cambridge University Press: Cambridge, UK.
13. Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. (2011). A quantum theoretical explanation for probability judgment errors. Psychological Review, 118, 193-218.
14. Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision making. Journal of Mathematical Psychology, 50, 220-241.
15. Bush, R. R. & Mosteller, F. (1951). A model for stimulus generalization and discrimination. Psychological Review, 58, 413-423.
16. Casale, M. B., Roeder, J. L., & Ashby, F. B. (2012). Analogical transfer in perceptual categorization. Memory & Cognition, 40, 434-449.
17. De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., & Storms, G. (2008). Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. Behavior Research Methods, 40, 1030-1048.
18. Eisler, H. & Ekman, G. (1959). A mechanism of subjective similarity. Acta Psychologica, 16, 1-10.
19. Evers, E. R. K. & Lakens, D. (2014). Revisiting Tversky's diagnosticity principle. Frontiers in Psychology, Article 875.
20. Grdenfors, P. (2000). Conceptual spaces: the geometry of thought. MIT Press.
21. Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. Cognitive Science, 7, 155-170.
22. Gleitman, L. R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. Cognition, 58, 321-376.
23. Goldstone, R. L. (1994). The role of similarity in categorization: providing a groundwork. Cognition, 52, 125-157.
24. Goodman, N. (1972). Seven structures on similarity. In N. Goodman, Problems and projects (pp. 437-447). Indianapolis: Bobbs-Merrill.
25. Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. Trends in Cognitive Sciences, 14, 357-364.
26. Hughes, R. I. G. (1989). The structure and interpretation of quantum mechanics. Harvard University press.
27. Hampton, J. A. (2007). Typicality, graded membership, and vagueness. Cognitive Science, 31, 355-384.
28. Isham, C. J. (1989). Lectures on quantum theory. Singapore: World Scientific.
29. Kintsch, W. (2014). Similarity as a function of semantic distance and amount of knowledge. Psychological Review, 121, 559-561.
30. 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.
31. Krumhansl, C. L. (1988). Testing the density hypothesis: comment on Corter.

Journal of Experimental Psychology: General, 117, 101-104.
32. Lambert-Mogiliansky, A., Zamir, S., & Zwirn, H. (2009). Type indeterminacy: A model of the KT (Kahneman-Tversky) man. Journal of Mathematical Psychology, 53, 349-361.
33. Larkey, L. B. & Love, B. C. (2003). CAB: Connectionist analogy builder. Cognitive Science, 27, 781-794.
34. Medin. D.L., Goldstone, R.L., & Gentner, D. (1993). Respects for Similarity. Psychological Review, 100, 254-278.
35. Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 104- 114.
36. Nosofsky, R. M. (1991). Stimulus bias, asymmetric similarity, and classification. Cognitive Psychology, 23, 94-140.
37. Oaksford, M. & Chater, N. (2009). Précis of Bayesian rationality: the probabilistic approach to human reasoning. Behavioral and Brain Sciences, 32, 69-120.
38. Parducci, A. (1965). Category judgment: A range-frequency model. Psychological Review, 72, 407-418.
39. Polk, T. A., Behensky, C., Gonzalez, R., & Smith, E. E. (2002). Rating the similarity of simple perceptual stimuli: asymmetries induced by manipulating exposure frequency. Cognition, 82, B75-B88.
40. Pothos, E. M. & Trueblood, J. S. (2015). Structured representations in a quantum probability model of similarity. Journal Mathematical Psychology, 64, 35-43.
41. Pothos, E. M., Busemeyer, J. R., & Trueblood, J. S. (2013). A quantum geometric model of similarity. Psychological Review, 120, 679-696.
42. Pothos, E. M., Yearsley, J., Barque-Duran, A Hampton, J. A., Busemeyer, J., Trueblood, J., (2015) Progress and current challenges with the Quantum Similarity Model. Frontiers in Psychology, 6:205. doi: 10.3389/fpsyg.2015.00205.
43. Pothos, E. M. (2005). The rules versus similarity distinction. Behavioral & Brain Sciences, 28, 1-49.
44. Pothos, E. M. & Busemeyer, J. R. (2009). A quantum probability explanation for violations of rational decision theory. Proceedings of the Royal Society B, 276, 2171-2178.
45. Pothos, E. M. & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? Behavioral & Brain Sciences, 36, 255-327.
46. Rosch, E. (1975). Cognitive reference points. Cognitive Psychology, 7, 532-547.
47. Shafir, E. B., Smith, E. E., & Osherson, D. N. (1990). Typicality and reasoning fallacies. Memory & Cognition, 18, 229-239.
48. Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. Science, 237, 1317-1323.
49. Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist networks. Artificial Intelligence, 46, 159-216.

50. Sloman, S. A. (1993). Feature-based induction. Cognitive Psychology, 25, 231-280.
51. Sloman, S. A. & Rips, L. J. (1998). Similarity as an explanatory construct. Cognition, 65, 87-101.
52. Tenenbaum, J. B, Kemp, C., Griffiths, T. L., & Goodman, N. (2011). How to grow a mind: statistics, structure, and abstraction. Science, 331, 1279-1285.
53. Trueblood, J. S. & Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. Cognitive Science, 35, 1518-1552.
54. Tversky, A. (1977). Features of Similarity. Psychological Review, 84, 327-352.
55. Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjuctive fallacy in probability judgment. Psychological Review, 90, 293-315.
56. Voorspoels, W., Vanpaemel, W., & Storms, G., (2008). Exemplar and prototypes in natural language categories: a typicality-based evaluation. Psychonomic Bulletin & Review, 15, 630-637,
57. White, L., Pothos, E., Busemeyer, J. (2014). Sometimes it does hurt to ask: the constructive role of articulating impressions. Cognition, 133, 48-64.
58. Yearsley, J. M., Pothos, E. M., Hampton, J. A., & Barque-Duran, A. (2014, in press). Towards a quantum probability theory of similarity judgments. Proceedings of the Quantum Interaction Conference.
59. Yukalov, V., & Sornette, D. (2010). Decision theory with prospect interference and entanglement. Theory and Decision, 70, 283-328.