Abstract
The transformational approach to similarity views similarity as a function of the complexity of the transformations required to ‘distort’ the representation of one object into another. These transformations may be more complex in one direction than the other, thus giving rise to asymmetric similarity relationships. Using the same-different paradigm we show that response times significantly differ depending on the direction of comparison, in line with the predictions made by the transformational approach. We discuss the implications of this result in reference to featural and spatial models.

Keywords: Similarity; asymmetry; transformations; representation; structure.

Introduction
Asymmetry is arguably the most counter-intuitive phenomenon in the study of similarity. Tversky (1977) was one of the first to argue that similarity is an asymmetric relation. In Tversky’s classic example, participants judged the similarity of North Korea to China to be greater than the similarity of China to North Korea. Asymmetries appear to be robust; evidence of asymmetric similarity has accrued across different stimuli (countries, geometric shapes, narratives, self concepts & music), measures (confusability vs. ratings) and species (e.g., non-human primates) (for previous evidence see Bartlett & Downing, 1988; Bowdle & Gentner, 1997; Catrambone, Beike & Niedenthal, 1996; Op de Beeck, Wagemans & Vogels, 2003; Tversky, 1977; but see also Gleitman et al, 1996). Given their pervasiveness, it is imperative that a model of similarity be able to capture asymmetries.

Initially, the similarity between two objects was considered to be the epitome of symmetry. The spatial model (Shepard, 1957) embodies symmetry in its fundamental axioms (i.e., the distance between two objects in a coordinate space is the same regardless of direction). It was because of this fundamental assumption that asymmetries gained theoretical attention. Tversky (1977) specifically put forward asymmetries as evidence against spatial models, and in favor of his own featural approach, the Contrast Model. Tversky argued that judgments of similarity could not be removed from the actual statements which formed the basis of them. The statement “A is like B” is unique in that it has a directional component; it has a referent (or base) B and a subject (or target) A, and the choice of object for these respective roles is unlikely to be arbitrary.

Generally, Tversky noticed that participants preferred the direction where similarity was maximized, and this involved selecting the most salient or prototypical object as the base object, as opposed to the target. Referring, once again, to the classic example, “North Korea is similar to Red China” is preferred to “North Korea is similar to Red China”. Tversky’s contrast model,

\[ \text{SIM}(A,B) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A) \]

defines the similarity of representations A and B as a function of their shared features, minus those distinctive to A, minus again those specific to B. The parameters \( \alpha, \beta, \) and \( \theta \) are weighting terms that depend on the task. In a non-directional judgment (“how similar are A and B?”) the distinctive feature sets of both objects are given equal weight (i.e., \( \alpha = \beta \)). In this case, similarity will necessarily be symmetric. However, asymmetries will arise when objects are subject to a directional comparison, in which case the distinctive features of one object may be weighted more heavily than those of the other (i.e., \( \alpha > \beta \)). This will give rise to asymmetries whenever the objects differ in salience or complexity, that is, \( f(A) \neq f(B) \).

There is an interesting consequence of this assumption in the Contrast Model, namely that if two objects differ in salience or complexity, then their self-similarity will also necessarily differ. In distance-based models of similarity, the similarity of an object to itself is the same, regardless of the object. In the Contrast Model, however, similarity, and with it self-similarity, has no inherent upper bound. In support, Tversky provides some evidence for differences in what might be construed as self-similarity in both ratings tasks (Gati & Tversky, 1982) and in the percentage of correct same responses on a same-different task (Rothkopf,
In response to Tversky’s critique of spatial models, Krumhansl (1978) put forward the distance-density model, an amended spatial model that allows asymmetries. Nosofsky (1991), then, demonstrated how asymmetries could be accounted for within a spatial approach to similarity through the general notion of a differential stimulus bias. Such a bias is associated with individual stimuli, and captures the fact that stimuli can differ in their salience or in the ease with which they are encoded. Nosofsky then went on to show that Krumhansl’s (1978) model, but also the Contrast Model (at least in specific versions) are all instances of the same general, stimulus bias framework.

What is common to all specific manifestations of the stimulus bias framework is that asymmetries arise as a consequence of the inherent properties of individual stimuli; they do not stem from the nature of the comparison process itself. This is in marked contrast to the transformational approach to similarity. As we will show below, it accommodates asymmetries based on stimulus salience or complexity, but also allows asymmetries in cases where no differences in complexity or salience exist.

The transformational approach, as proposed by Hahn et al. (2003), assumes that similarity is determined by the complexity of ‘transforming’ or ‘distorting’ the representation of one object in another. Similar objects will require simple transformations, dissimilar objects will require complex ones. Several studies have provided support for a transformational approach. In their original paper, Hahn et al. (2003) demonstrated that transformation distance - operationalized as the number of individual operations required - predicted similarity ratings across 3 experiments using a range of materials (dot patterns, simple geometric shapes and Lego bricks). Furthermore, a simple featural model fared poorly in comparison.

Closely related to the aims of the present investigation, Hahn, Close and Graf (2009) exploited directional similarity judgments (‘how similar is A to B?’) to test the transformational account. Transformational complexity can differ readily depending on direction: spilling water from a cup, for example, is easier than gathering the spilled water back in. Any such directional difference should give rise to attendant differences in perceived similarity, and hence asymmetric similarity between the two comparison points. Hahn et al. (2009) tested whether an inherent sense of direction could be artificially induced. To this end, they showed participants short animations of one familiar basic level object morphing into another. After viewing the animation, participants rated the similarity of objects drawn from the morph continuum. Directional similarity ratings for the exact same comparisons were higher when the referent object (or base) had appeared first in the preceding animation; that is, ratings were higher when the direction of the similarity comparison matched the direction of the preceding animation.

Given that the experimental manipulation involved only the direction of the preceding animation, it is hard to see how these results could be explained through differential salience or complexity of the two objects being compared. Instead, it seems that perception of the ease or naturalness of the transformation itself was being affected. In other words, it seems that the directional asymmetries that arose here stemmed from the nature of the comparison process itself, not from intrinsic properties of the individual stimuli.

The current paper seeks to examine further the kinds of asymmetries that can arise on the transformational account, using both a different stimulus domain, and an implicit measure of similarity, as opposed to explicit ratings.

**Experiment**

In this study, participants had to compare two pairs of shapes that varied across two dimensions (shape and color), as illustrated in Figure 1. Stimuli of this kind have been used extensively to study feature binding (Cheries et al., 2006; Kaldy & Leslie, 2003). More recently, they have also been used to test structure-based models of similarity (Larkey & Markman, 2005).

Hodgetts et al. (2009) provided a simple coding language for this domain, and found excellent quantitative predictions between the transformational predictions derived from this coding language and perceived similarity in both a 2-alternative forced-choice task and a direct ratings task. Moreover, the data fits were superior to a number of structural alignment based models (SIAM, Goldstone, 1994; SME, Gentner, 1989).

As the present test stimuli are a subset of those used by Hodgetts et al., we will describe these materials and the associated transformations in detail. On each of the two dimensions (i.e., color and shape) there are 14 possible structural relationships across the two composite ‘objects’. As color and shape variation can furthermore be factorially combined, there are 196 different quadruples in this domain. Hodgetts et al. (2009) posited three general transformations, or operations, for this domain. These are applied to the base pair (‘object 1’) in order to modify it so as to generate the target pair (‘object 2’). These transformation operations are:

1. **Create feature** – taking the base pair we apply this operation to create a new feature that is unique to the target pair.
2. **Apply feature** – this operation takes an object or entity that is currently available (by being present in the base or by having been created via step (1)
and applies it to one or both of the objects in the target pair.

3) **Swap** – this swaps features between a pair of objects or swaps the object in its entirety (i.e., on both dimensions).

Figure 2 demonstrates a transformational sequence as determined by this coding scheme.

![Figure 2: An example of a transformational sequence being applied (distance = 3).](image)

In a non-directional context, this coding scheme takes the distance associated with the greatest complexity, the MAX-distance between the two pairs, as its overall prediction. However, when a comparison is directional, the distance is simply the number of steps from base to target. Out of the 196 possible comparisons, 122 are asymmetrical; here, the transformation distance between the two pairs depends on the direction, that is, which object pair is chosen as the base from which the other is derived. An example is depicted in Figure 3.

As can be observed, the transformation distance left to right is 2, whereas it is 4 in the opposite direction. In complexity terms, the former transformation is simpler and therefore associated with greater perceived similarity. In terms of the coding scheme, the code is shorter left-to-right because ‘applying’ a feature to two shapes is as complex as applying it to one, that is:

\[
\text{create (circle) + apply (circle)} = 2
\]

By contrast, in the case of the longer code, both shapes must be created and applied separately; there are no concessions for applying two objects because the shapes differ:

\[
\text{create (square) + create (triangle) + apply (square) + apply (triangle)} = 4
\]

Because the differences in complexity give rise to differences in code length (that is, the specification of the instructions required to perform the transformation) we will simply refer to the two directions as ‘long’ and ‘short’ in the following, where stimuli of varying transformation distance are considered.

It is worth noting that the transformational coding scheme of Hodgetts et al. was not derived with asymmetries explicitly in mind, rather they are simply a consequence of the operations associated with these specific stimuli. Moreover, simply considering these shapes, it seems difficult to assign differential salience or complexity. Under the Contrast Model, the right-hand object could be argued to possess greater ‘goodness of form’ and therefore be more salient. Alternatively, the left object, by possessing a greater number of unique features could be argued to be more complex, and therefore more salient. In other words, it seems difficult, from the perspective of the Contrast Model to make *a priori* predictions whether or not these stimuli should give rise to asymmetries, and, if yes, what the direction of the asymmetry should be.

To test experimentally the asymmetry predictions of the transformational account we required a directional task. We chose a speeded same-different (perceptual matching) task for this purpose. In this task, participants are briefly presented with two stimuli, one after the other, and asked whether the second is the same as the first. This task is naturally directional, in that it involves comparing the second stimulus to the first. Hence, we assume that the first object represents the base of the comparison, and the second object the target; that is, the sequential presentation corresponds to the directional comparison ‘how similar is second object’ to ‘first object’?’. This task provides a paradigm for measuring similarity implicitly, because response times on this task have been found to depend lawfully on the degree of similarity between items. Specifically, participants take longer to correctly identify two stimuli as different when they are more similar (e.g., Cohen & Nosofsky, 2000). Consequently, if the predictions of our transformational account are correct, participants should take longer to correctly respond ‘different’ if the transformational relationship between the first and second stimulus is simple, than if it is complex.

![Figure 3: An example of an asymmetric relationship.](image)

**Participants** 39 psychology undergraduates.

**Materials** The task was presented on a 19” LCD monitor with a screen refresh rate of 60 Hz. ‘Different’ trials consisted of two sequential presentations of object pairs that varied in color and shape. A sequential presentation allowed directional predictions to be tested directly. As code lengths \(a\) to \(b\) and \(b\) to \(a\) vary for each stimulus group in our set (see Figure 3), we could present both directions for all of our
comparison stimuli and simply contrast the response times for the direction of the shortest code length with those of the longest code length across all stimuli. The shapes were created using the AutoShape function on Microsoft Publisher. There were three possible features on each dimension (shape = triangle, square circle; color = yellow, purple, green; for a detailed description see Hodgetts et al. (2009)). Each shape was 2.5 cm wide x 2.5 cm tall. Shapes within a pair were separated by a horizontal distance of 0.5 cm. The screen location of pairs on a given trial was determined by randomly combining set values on each screen axis (i.e., 10, 20, 30, 40, 50, 60, 70, 80 and 90). The stimulus duration for a given pair was 833 ms (50 frames) with an ISI of 17 ms (1 frame). A response could be given at the onset of the second stimulus.

From the 196 comparisons available to Hodgetts et al. (2009), 122 contained directional differences in transformation complexity. Participants in each group received 244 different trials and 244 same trials across two blocks. To manipulate direction, participants were shown 61 comparisons base to target and target to base in each block. Participants responded same or different by pressing the appropriate key (‘Z’ or ‘M’). Participants could respond at the onset of the second stimulus. No response deadline was imposed but participants were urged to respond as quickly as possible.

Results
For analysis, we first compared mean reaction times on correct ‘different’ trials for each of the tested directions (short and long). Reaction times (RT) more than three standard deviations above and below the overall mean were removed from analysis. Two participants were removed for low overall accuracy (<50%). As explained above, the transformational approach predicts slower response times in the direction of the shorter or simpler code, as this corresponds to greater similarity. The graph in Figure 4 confirms this predicted pattern of results. The slower observed RT in the ‘short’ direction indicates greater perceived similarity (mean = 344.9 ms; SE = 24.8). Correspondingly, the faster RT in direction of the long transformation indicates that these comparisons are less similar (mean = 633 ms; SE = 26.8). A within-subjects t-test, shows this difference to be statistically significant (t (1, 36) = 2.5, p < .05). The specific objects compared in each case are identical; the only difference is the order of their presentation. Crucially, this order differentially affects the complexity of the transformation that manipulates the object representations.

Self-similarity and complexity
As noted above, it is unclear what the predictions of the Contrast Model are for these materials. However, it is possible to test, after the fact, whether the differences we found are compatible with the model, by investigating the ‘same’-trials. As noted in the introduction the Contrast Model assumes that if object A is more complex or otherwise more salient than object B, that is if f(A) > f(B), then “B is like A” will be preferred to the opposite. At the same time, if f(A) > f(B) then the self-similarity of object A will be greater than that of object B. That is, the more salient or more complex object will always be more similar to itself than the less salient object.

To assess ‘self-similarity’, we compared reaction times for correct same trials as an index of self-similarity. If differential salience/complexity, as assumed by the Contrast Model, can explain these results, then reaction times should be greater for those objects that form the base objects in the comparison direction with the ‘short’ code (i.e., object a in Figure 1 will have a slower RT when compared to b). The graph in Figure 5 shows mean reaction times for the base and target objects in the short condition. The graph shows slower response times, on average, for the base objects when compared with the target objects for the short direction. This relationship was also born out statistically; a within samples t-test yielded a significant difference (t (1, 113) = 4.5, p < 0.01). The longer response times on the same trials for the base pairs indicate that the result, while not
clearly predicted by the Contrast Model, is at least compatible with it.

**Discussion**

The fact that these asymmetries are shown, even with reaction time, is evidence that asymmetries are not only relevant in explicit ratings task but also in an online task that requires rapid perceptual matching between objects.

Furthermore, the results from this experiment support the idea that transformation distance can act as a useful metric in predicting asymmetric similarity between objects. This, in turn, supports not only the general idea that similarity can be conceptualized by transformational relationships, but also that the specific coding language, devised to reflect the representations of these objects, is making the right psychological predictions. This result extends on the work of Hahn et al. (2009), by showing transformation based asymmetries in a different domain, and with an implicit measure of similarity, as opposed to explicit ratings.

On the transformational account, asymmetries arise when one direction is simply easier than the other in terms of transformational complexity or code length (i.e., requires fewer instructions to transform). In the current experiment, asymmetric similarities are manifest in the longer response times that exist in the direction where transformation distance is less.

Moreover, unlike the Contrast Model, our coding scheme makes unequivocal predictions both about the existence of asymmetries with these objects, and about their direction. Nevertheless, in this experiment, the results are compatible with both the transformational account and the Contrast Model. Analysis of the ‘same’-trials showed the preferred base objects to possess greater self-similarity. According to the contrast model, asymmetries arise when object A is more salient (or complex) than object B, meaning the statement “B is like A” will be preferred over “A is like B”. From this, the contrast model predicts greater self-similarity for the salient (or preferred base) object, which is measured here by response time on ‘same’ trials. Slower responses for the preferred base objects support the idea that they possess greater salience than the preferred target. Whilst this does not negate a transformational explanation, it necessarily does not refute the Contrast Model in this context: asymmetries can be argued to have emerged from the differential salience of the compared objects.

In this regard, the present results complement those of Hahn et al. (2009), by demonstrating how the transformational framework applies to asymmetries arising from differential object complexity. Although it is the complexity of the transformations relating the two objects, on this account, not the complexity of the objects themselves, there will typically be a systematic connection between the two. The fact that one of the ‘objects’ in Figure 3 above contains two different shapes, whereas the other ‘object’ contains only 1, has knock-on effects for the transformations that relate them. “Applying” two different features costs more than applying the same feature twice. Critically, these predictions come about naturally without additional parameters. The same is true of other potential differences in complexity; for example, the Contrast Model also predicts that adding distinctive features to the base will increase the magnitude of the asymmetry. An attendant complexity difference in the associated transformation arises in this case because deleting features requires a less complete specification of those features than inserting features, leading to a shorter code overall (see also Hahn & Bailey, 2005, for evidence to this effect in the domain of word similarity).

Interestingly, this domain contains a further possibility for testing not only the Contrast Model, but the entire general, differential bias framework. Nosofsky (1991) highlights a simple way in which the differential bias hypothesis could be falsified. Any additive similarity and bias model implies the following transitivity condition: If \( p(i, j) \geq p(j, i) \), and \( p(j, k) \geq p(k, j) \), then \( p(i, k) \geq p(k, i) \). Essentially, if an asymmetry exists for \( i \) and \( j \) then at least one asymmetry must exist for \( i \) and \( k \) or \( j \) and \( k \). If one conceptualizes these objects as possessing differential biases, outside the comparison, then one can see how asymmetries must be transitive in these triple scenarios. Identifying and demonstrating violations of this transitivity condition would be of enormous theoretical importance in terms of modeling asymmetries. The coding scheme presented here, however, allows such ‘isolated asymmetries’. Hence testing the relevant triplets seems a priority for future research.

Finally, the most general conceptual difference between the transformational account and both spatial and featural models is that the transformational account allows for asymmetries within a structural framework. Whereas the two traditional models assume very simple and specific representations, that is, features sets or separable continuous dimensions, the transformational account is one, out of a number of recent accounts, that is applicable to structured representations. Structure refers not only to the features that make up an object but also to the relations between these features (Biederman, 1985; Gentner, 1983, 1989; Hahn, Chater & Richardson, 2003; Markman and Gentner, 1993a, 1993b). For the objects in the domain examined here, certain transformations, such as the swap transformation, implicitly suggest that the left-of/right-of relations between objects are represented and thus manipulated via the swap. While structure-based transformations do not govern each comparison, it does allow for a level of complexity not permitted under any simple featural or spatial model.

Crucially, there is little previous evidence relating structural models of similarity and asymmetry. The only previous study was conducted by Bowdle and Gentner (1997), who combined structure mapping theory with Grice’s (1975) pragmatic principle of informativity. Structure mapping theory states that the similarity between
two objects is calculated by structurally aligning object representations (Gentner, 1983; 1989; Markman & Gentner, 1993a). Generally, they argue that asymmetries will occur when the base is more systematic than the target as it then 'lends' structure accordingly. Similarly to the transformational account, this model requires alignment between objects and so remains in the frame of comparison. However, conceptualizing the current results in terms of systematicity or informativity is difficult and seems far less intuitive than it does in the traditional domain of structure mapping models which involve items such as short narratives). In other words, it seems that models within the structural alignment framework do not predict asymmetries for our items.

In summary, we have provided evidence that asymmetries in directional similarity comparisons can be accurately predicted by a difference in transformational complexity. Furthermore this accuracy is demonstrated in an implicit speeded task. Whilst these results are at least compatible with Contrast Model, there are further tests that could be applied within this domain that could test further the adequacy of the differential bias framework. Hence, this domain recommends itself for further exploration of asymmetries as a diagnostic test for models of similarity.

Acknowledgments

This research was supported by European Commission grant 51652 (NEST). Nick Chater was supported by a Leverhulme Major Research Fellowship. We would like to thank members of the School of Psychology for their constructive feedback and support.

References