Towards a General Law of Numerical/Object Identity

Felice L. Bedford

University of Arizona, Tucson, USA[*]

---

[*] Correspondence should be sent to Felice L. Bedford, University of Arizona, 312 Psychology Building, Tucson, AZ 85721, USA (email: bedford@u.arizona.edu)

Abstract

A theory of object identity, knowing whether or not two glimpses refer to the exact same object, is presented in two parts. Part I argues that apparent motion, prism adaptation, ventriloquism, priming, stereopsis, and Gestalt grouping all require the identity decision. It is argued that object identity is general and required when samples come from different times, places, modalities, and eyes. Part II argues there is a common solution at a sufficient level of abstraction. Sample 1 and sample 2 are regarded as two forms which differ by a transformation corresponding to one of five geometries, Euclidean, Similarity, Affine, Projective, and Topology (Klein, 1893), that nest within each other like a set of Russian nesting dolls. Identity is resolved from the lowest level of the hierarchy available in the situation, producing a flexible solution whereby the same two samples will sometimes refer to the same object and sometimes not.

We believe the world contains objects and events that extend in space and endure in time. A general problem for observers is to extract those meaningful individuals. One specific problem is how we determine that a stimulus we are looking at now reflects the *exact same* item as one encountered before, an issue known as identity, numerical identity, or object identity (e.g. Hirsch, 1982; Leslie, Xu, Tremoulet, and Scholl, 1998; Meltzoff and Moore, 1998; Spelke, et al., 1995; Strawson, 1959; Xu and Carey, 1996).

Several challenges await an observer who must make an object identity decision. First, movement of the eyes, head, and body contribute to an ambiguous retinal image. As we move ourselves around, the image on the retina changes. A person you see is first large, then small, on your retina. How do you know you are looking at the same person? You turn sidewise while looking at the closet door and the image transforms from a rectangle to a trapezoid. How do you know it is the same door? Traditionally, the perceptual constancies have described the accomplishment that the object is perceived as staying constant despite the changing retinal image (Helmholtz, 1860; Hering, 1905; see Woodworth, 1938).

Second, the objects of our perceptions themselves can move and change (e.g. Shepard, 1988). Rocks can be thrown, sponges squished, leaves on a branch blow in a breeze, tigers pounce, people can twist and turn, and frogs can turn into princes, at least once upon a time. The image cannot be the same before and after each event, yet the pre and post-transformed stimuli often refer to the same object. When an object changes, we need both to perceive those changes to the object as well as know we are still dealing with the same object.

Third, identifying enduring objects is further complicated by a lack of continuity of sensory stimulation (e.g. Michotte, 1963; Spelke and Kestenbaum, 1986; Xu and Carey, 1996). If you see Jack's beanstalk grow before your very eyes, you accept it is the same object despite the unprecedented growth. But stimuli go away and reappear when we blink our eyes, turn away for an instant, attend for a while to something else, or come back later in the day. We do not always have continuity to tell us that a transformation, however bizarre - or ordinary - nonetheless reflects the same object.

While these three problems for object identity are already general and formidable, progress is sometimes obtained, and solutions simplified, by becoming more general still (see Shepard, 1994, 2001). This article makes two claims: 1) Object identity is universal and abstract. Not only is there a problem across time, as it is usually formulated, but also across modalities, spatial locations, and even the left and right eyes. Asked generally, how does the observer determine when any two non-identical samples arise from the same object? Because much of interacting with the world involves non-identical samples, I claim the same object identity decision plays a role in nearly all domains. 2) There are a core set of criteria abstract enough to apply to all object identity domains. These apply wherever the samples originate (time, space,

modality, or eye), whatever the scale (short or long time frame, or distance), the level of the computation (preconstancy/retina and postconstancy/postretina) or the size of the sample (point sources, extended contours). We believe this core can be described by geometry - not just the familiar Euclidean geometry, but a whole family of at least 5 different-sized geometries (Klein, 1893/1957) that nest within one another like a set of Russian nesting dolls. The probability that two samples will be judged to refer to one object decreases as the smallest geometry within which the they are equivalent gets larger.

Part I defends the first claim. It is shown how a number of classic and popular phenomena all depend critically on the object identity decision. Within many of those domains, a principle equivalent to object identity has been independently rediscovered, but the areas rarely interact with one another and are considered to reflect independent modules of mind Part II defends the second claim. It is shown how the mathematician Klein's hierarchy of geometries is essential for object identity and how it can be applied to the diverse manifestations of identity claimed in Part I, even when they do not appear to involve shape or geometry. Discussion follows each part.

Part I: The same problem resurfaces in different modules

For each phenomenon, a) the phenomenon is briefly described, b) a case is made for object identity, c) the role that object identity has played in the literature is noted, d) any criteria of identity that have been uncovered also noted, and finally d) the relevance of the laboratory effect for perception outside the laboratory is considered. These and other effects which

arguably require a same-object judgement to occur are summarized in Table 1.

1.1 Apparent motion.

Apparent motion refers to the well-known illusory experience of motion when all the stimuli are actually stationary. A stimulus is shown in one location for a brief period of time and is replaced by a second stimulus in a different location a short time later. The sequence is repeated a number of times and observers see a stimulus moving back and forth between the two locations.

Apparent motion is also the prototypic laboratory phenomenon that involves object identity. Rock (1983) and Shepard (1988) argue that an interpretation of a single object moving back and forth between the two locations - even with intermediate positions missing - is preferable to accepting the coincidence that two identical or similar looking objects are appearing and disappearing in alternation (see also Ullman, 1979). If the two samples are judged to refer to one object, one object in motion is experienced. If the two samples are judged to refer to two objects, then two objects flashing on and off are seen and there is no impression of motion. It is generally accepted that the phenomenon rests on misidentifying the two stimuli as two different glimpses of the same object at two different times, but see Berbaum et al. (1981) for an exception.

A demonstration that object identity is involved is Sigman and Rock's (1974; see also Rock, 1983) variation with a screen added to the two stimuli. At time one, there is a small black dot in position A and a big screen covering position B, and at time two, the screen covers position A, and there is a small black dot in position B. The display looks as if there is a screen moving back and forth to successively occlude and reveal two dots, one on each side. The addition of the

screen invites a different explanation of the mysterious sensory appearance and disappearance of two stimuli: rather than one object moving back and forth, a screen is moving back and forth in front of two objects.

Although object identity is recognized as critical for apparent motion, the criteria of object identity in apparent motion - the rules by which the observer determines whether the two samples refer to the same object - are unfortunately not always clear. Only a few investigators have explicitly used the phenomenon as a paradigm to uncover the

Same-object phenomena	Brief Description	S	T	M	E	Comments
Apparent Motion	(Mis)perception of an object moving from one location to another	2	2	1	1	
Apparent Transformation	(Mis)perception of an object transforming	1	2	1	1	not a recognized class but should be. Apparent motion is a subset.
Priming	Facilitation in the processing of a stimulus as a result of prior presentation of a stimulus	1	2	1	1	
Stereopsis	Perception of distance of a stimulus after matching left and right eye stimuli	1 or 2	1	1	2	if L and R stimuli fall on corresponding points, then 1 spatial location; otherwise 2.
Gestalt grouping	Experience that parts of a spatially distributed display spontaneously group together	2	1	1	1	A collection of different phenomena unified by same object "judgment"
Ventriloquism Effect	(Mis) perception of sound as coming from visual position	2	var	2	1	Samples usually extend in time; if discontinuous, then > 1 time sample; otherwise 1.
Visual or auditory capture	Immediate predominance of one modality over another when there is a conflict	1 or 2	1 or 2	2	1	e.g. visual capture of proprio. for size (S=1 T=1); visual capture of proprio. For location (S=2 T=1); auditory capture of vision for temporal sequences (S=1 T=2). ventriloquism is a subset.
Prism adaptation	Recalibration of proprioception and/or vision for spatial location	2	var	2	1	Exposure to prism can be one long "trial" (1 time sample) or multiple discontinuous trials (>1 time samples)
Cross-modal adaptation	Recalibration of a modality following conflicts of orientation or size etc.	1	var	2	1	Prism adaptation is a subset
Perceptual constancy	Perception that properties (e.g. size, shape) remain constant despite varying retinal images	1 or 2	1 or 2	1	1	See text-Part II
McCollough Effect	Alternating red-vertical and green-horizontal grids leads to red/green color contingent on retinal orientation	1	2	1	1	Grids "judged" to be 2 glimpses of same object to get effect. (see Bedford, 1995)
Contingent aftereffects	General set of phenomena where selected pairings of stimuli leads to one proximal dimension becoming contingent on another	1	2	1	1	Same as above
McGurk Effect	Auditory phoneme perception altered by simultaneous visual display of a person uttering a	1	1	2	1	must auditory and visual speech be judged at some level to originate from the same speaker?

	different phoneme					
Pavlovian conditioning	Association of two discrete samples judged to originate from the same source	2	2	2	1	must samples be "judged" to originate from the same source? samples refer to standard successive conditioning (e.g. tone-shock; taste aversions)

Table 1. Object identity phenomena. A list of effects that appear to require the "same-object" decision. Following a description of each effect is an indication of the number of discrete samples from space (S), time (T), modality (M), and eye (E) that are involved.

Pavlovian conditioning	Association of two discrete samples judged to originate from the same source	2	2	2	1	

criteria (e.g. Warren, 1977, Chen, 1982). In addition, dependent measures in apparent motion have varied considerably, including asking subjects directly whether they experience one object or two (Warren, 1977), asking subjects to describe any motion they see (Berbaum et al., 1981), obtaining a measure of whether the experienced motion is "good" (Dawson, 1986), and what direction of motion is seen when there is a choice (Chen, 1982); perhaps as a consequence, results are often contradictory (see e.g. Mack, Klein, Hill and Palumbo, 1989). Changing shape from one stimulus to the next hurts apparent motion (Orlansky, 1940; Mack et al. 1989) but others argue it is largely irrelevant (Kolers and Pomerantz, 1971; Kolers, 1972; Navon, 1976); changing orientation influences apparent motion for some (Green, 1986; Ullman, 1980), but not for others (Burt and Sperling, 1981); size changes are influential (Navon, 1976) but they are also not (Burt and Sperling, 1981); rigid motion is critical (Farrell and Shepard, 1981; Shepard, 1982; Shepard and Judd, 1976; Warren, 1977) but non-rigid plastic deformations are seen (Kolers and Pomerantz, 1971; Shepard and Judd, 1976; Farrell and Shepard, 1981). These results will be explained in Part II. For now, note that regardless of the criteria by which object identity is determined, the fact that object identity is necessary to produce apparent motion is clear.

Finally, note also that apparent motion is a reflection of everyday perception. The integration that occurs simulates "real world" circumstances where there are temporal gaps in perceiving. During saccades, when the eyes dart from one location to another, there is saccadic suppression. Consequently, three times per second there is effectively no visual information on the retina which creates temporally discontinuous samples. Gaps that occur during eye blinking approximately once per second may also be in the same time frame as apparent motion. For everyday perception, it is therefore critical that the observer determine whether these different samples in microtime refer to the same object in order to accurately parse the world into separate enduring objects and to track those objects if they move.

1.2 Prism Adaptation.

The object identity decision is also critical for "prism adaptation" (Welch, 1972), an unrelated phenomenon that rarely makes contact with apparent motion or object perception, and rarely recognizes the importance of identity.

Prism adaptation is an example of cross-modal perceptual learning made popular by Stratton's work in 1897. To produce adaptation, subjects look through prisms or lenses which distort vision. Initially, reaching, pointing, and maneuvering in the world are predictably disrupted but following practice, effective coordination is restored. Genuine adaptation is not due to an intentional compensation for the distortion, but to a change in the underlying perception of locations (see e.g. Bedford, 1993b; Harris, 1965; Redding and Wallace, 1976). Stratton turned the world upside-down visually but most subsequent researchers have used a small lateral displacement for systematic investigations. When the observer sees his or her hand through the prism, the hand feels like it is one place, but looks like it is in another. It is this disagreement between vision and proprioception/touch which leads to the underlying perceptual shift in one modality or the other.

Turning to object identity, Welch (1972; see also 1978, 1986) has shown that whether or not adaptation occurs depends on whether the subject believes the differing visual and

touch signals come from the same arm. He devised a procedure whereby each subject was misled into believing that someone else's hand was actually his own. A confederate wore a glove that was covered in luminous paint in an otherwise dark room while the subject's glove was not painted. A prism experiment was simulated by having the confederate's hand moving in synchrony with a subject's hand as he attempted to point to targets, but shifted by a constant amount to the right, as it would be if the subject were looking at his own hand through a prism. To decrease the likelihood that subjects would catch on to the trickery, they were also told that any detected inaccuracies in pointing were due to a prism, when in fact they were looking only through window glass. The experiment led to genuine adaptation in the felt position of the hand, indistinguishable from the results of a "normal" prism experiment.

The belief that the differing signals from the visual position of the hand and the felt position of the hand were from the same hand led to adaptation, even though in this situation they were actually due to two different hands. In the same study, the reverse situation led to an attenuation of adaptation: When subjects looked through a prism but were mislead into believing the discrepant visual position of their hand was really an image of someone else's hand, adaptation was reduced. A disagreement between vision and touch was insufficient without the decision that the signals come from the same object. Welch called the belief that the discrepant visual and felt position signals come from the same hand under typical prismatic displacement the "unity assumption". His important discovery has been essentially ignored in adaptation research.

Another experimental demonstration of object identity in prism adaptation may come from an experiment by Hein, performed in the 1960's (Hein, personal communication, 1992). Each subject held a pencil in his hand which he used to aim for targets. The subject was allowed to see the pencil but not his own hand, which was hidden beneath a table. Although subjects looked through prisms while pointing, no adaptation occurred and the study was not published. Note, however, that while a subject could obtain the felt position of his hand, he could not see his hand, and while he could see the pencil, he could not clearly get the felt position of the pencil. The differing visual (v) and touch (t) signals may not have been judged to come from the same object, which precluded adaptation.

The experiments all show that disagreement between v and t *per se* does not lead to adaptation, but rather the disagreement along with a decision that v and t must come from the same object - the unity assumption or the object identity decision. Thus, object identity is an issue not just when samples come from different times, but also when they come from different modalities.

Why should a decision about *objects* be crucial for decisions about *space* in everyday perception? I have argued elsewhere (Bedford, 1993, 1999; see also 1993b, 1993c, 1995) that it is a key part of the solution to what appears to be a general paradox for perceptual plasticity. Perceptual mechanisms should change only in the event that an error has been detected. Otherwise, perception would be chaotic with ever-changing images, sounds and physical sensations. Perception needs to be stable and accurate in order to obtain reliable information from the world. This leads to the puzzle: If information about the world comes through the sensory systems, how can the sensory systems ever indicate that the sensory systems themselves aren't

functioning correctly? Why wouldn't any new sensory stimulation be interpreted as a something new occurring in the world?

The answer to the puzzle is that not all the information about the world does come through the senses; a priori internal constraints about the way the world works can be used to check the sensory information to determine if it is correct. The answer to why objects are relevant for adaptation and space perception is that many of those constraints center around objects. A constraint that an object cannot be in more than one place at the same time (Bedford, 1993; Narter, 1997; Rosser, Narter and Paullette, 1995; Wilcox, Nadel, and Rosser, 1996; Xu and Carey, 1996) along with a decision about object identity determines how to interpret a situation with different location values from vision and touch. Consider briefly three different possible reactions to the disagreement between vision and touch (Bedford, 1999). 1) The disagreement could be taken to indicate something new about the world: an object can be in two places at the same time; 2) the values could be different only because they are referring to different objects; 3) the different values could indicate that there is something wrong with the observer. Since one object cannot be in two places at the same time, if the sensory information suggests otherwise then there is a sensory error which should be fixed. In prism adaptation, the "one object - one place" constraint typically steers us away from the first alternative and a "same object" judgment away from the second. The object identity decision is critical in everyday perception (see also section 1.4) to determine if the disagreement between modalities is due to an internal error that must be corrected.

1.3 Ventriloquism Effect.

The object identity decision is also essential for producing the ventriloquism effect (Radeau and Bertelson, 1976). The ventriloquism effect, like prism adaptation, involves a difference between modalities about location, but the modalities are vision and audition rather than vision and touch, and the phenomenon refers to an immediate reaction rather than a long-term change following repeated experience. The phenomenon takes its name from the old vaudeville inspired entertainment where a performer, the ventriloquist, and his doll, the dummy "interact" with one another. When it is the dummy's "turn" to "talk", the ventriloquist talks without moving his lips and at the same time moves the dummy's mouth. The illusion is compelling: it sounds like the dummy is actually talking. That is, you mislocalize the sound of the voice as coming from the dummy rather than real speaker. Far from original theater claims that the ventriloquist is practiced in "throwing his voice", the illusion is created in the mind of the perceiver, not the performer. There is a disagreement between vision and audition: The visual stimulus (the moving mouth of the dummy) is in location x, but the auditory stimulus (the speech) is in location y. The illusion of ventriloquism is the misperception of the auditory stimulus as coming from location x rather than y. The correct auditory location, y, is registered by using the timing and intensity differences between the two ears, but by the time the information reaches conscious experience, it has shifted to x.

Radeau and Bertelson (1974; 1976) described the ventriloquism effect as a spatial discrepancy between visual and auditory signals which can be related to a *single source, event, or object*. Shortly thereafter (1977), they call this process "pairing" (see also Bertelson, 1996; Radeau,

1994)[1], after Epstein (1975) and Wallach (1968). You mislocalize the auditory location only because you have -in this case mistakenly -judged that the visual and auditory signals come from the *same* object, the dummy. While we are not easily misled, the carefully crafted display where the real speaker does not move his lips, the dummy's lips move in synch with the voice, the dummy's appearance is human-like, and the knowledge that humans talk, make this interpretation very likely. But if you had decided that the differing visual and auditory signals, x and y, come from *different* objects then there will be no ventriloquism effect.

I offer the following analysis of ventriloquism to make it plain why the identity decision is critical. Imagine whether you would get the ventriloquism effect if the dummy's lips weren't moving; how about if the dummy had no lips? No facial features? Surely not if the ventriloquist held a plain cantaloupe on his lap. This situation happens all the time; for instance, as you read this article, the TV is on in the background, out of sight. The visual signal you are receiving is in location "x" and the auditory signal in location "y". Yet you typically do not mislocalize the sound of the TV as coming from the book. This is because you have determined that they are two different objects. That is, the constraint that an object cannot be in more than one place at one time and associated arguments apply to ventriloquism as well as prism adaptation. If the differing visual and auditory signals are judged to refer to the same object, then

you know that your sensory systems have made a mistake. The mislocalization that characterizes the ventriloquism effect is an attempt to preserve the experience of a coherent world in spite of what is believed to be an error. The auditory system must have provided the wrong value, resulting in greater reliance on the visual value. But if the differing visual and auditory signals are judged to refer to two different objects, then everything is as it should be and there is no need to intervene. There is no constraint that *two objects* cannot be in two different places at one time.

The data on factors which influence the ventriloquism effect support the view that the object identity decision is crucial. Early experiments on the ventriloquism effect suggested that the realism of the displays is influential (e.g. Jackson, 1953; Jack and Thurlow, 1973; Thurlow and Jack, 1973). If the visual stimulus is a "whistling" tea kettle with steam, and the auditory stimulus is the sound of the whistle but is actually coming physically from a different location, a strong effect is obtained. But if the stimuli are replaced with lights and bells at the same spatial separation, the effect is reduced. Whistling tea kettles and the sound of a whistle are presumably more likely to be judged to refer the same object than items which have not previously been associated. A puppet and a voice track in a different location led to a compelling illusion, but this was reduced when the facial features were removed from the puppet. For a visual display of a person with his/her mouth moving, a synchronized voice in a different location led to capture of vision over audition 78% of the time; when a light of changing intensity was substituted for the person, the effect was reduced to 49%. There is also evidence that direct instructions to subjects about whether they will be experiencing one object or two

---

[1]The process of pairing was initially (Radeau and Bertelson, 1977) intended to refer to adaptation following prolonged contact with visual-auditory discrepancy, distinguished from "fusion", which they used to refer to the immediate experience that occurs in the ventriloquism situation. In their more recent works, pairing was extended to encompass the short-term as well as the long term reaction to the discrepancy.

objects influences how compelling the illusion is. Subjects who are not made to attend to the fact that the auditory and visual signals may come from two different objects are more likely to show an effect (e.g. Warren, Welch, and McCarthy, 1981).

Finally, synchrony of the visual and auditory signals is critical; remove it, and the ventriloquism effect is reduced, both when the stimuli are realistic (Radeau and Bertelson, 1977) and when they are flashes of light and sounds devoid of meaning (Choe, Welch, Gilford, and Joula, 1975; Radeau and Bertelson, 1987; Thomas, 1941). Radeau and Bertelson argue that synchrony is an important determinant of pairing and suggest it is reminiscent of the Gestalt principle of common fate. Presumably, if the onsets and offsets of the lights and tones are made different, one has logically reduced the possibility that the lights and sounds arise from the same object. If one object is responsible for both signals, then those signals may appear and disappear at the same times. But if lights and sounds are independently generated, then they are likely to come from completely different objects and events. Thus synchrony should be a powerful cue as to whether visual and auditory information arise from the same source, and it appears to be.

The factors that influence the ventriloquism effect also elucidate the criteria of object identity. Radeau and Bertelson (1974) explicitly suggested that "It should be possible to specify the conditions for single event interpretation, though the present study was not designed to analyze them", and later (1977, 1987) set out to do so. They divide factors into structural, or data driven, or bottom-up, properties such as synchrony on the one hand, and cognitive factors, which include instructions and realism (see also "compellingness", Welch and Warren, 1980; 1986) on the other.

Recently, Radeau (1994) has claimed that the only genuine influence on the ventriloquism effect are from the bottom-up factors, despite the contradictory evidence from other studies that top-down properties, such as verbal instructions, are influential. This debate is currently unresolved. However, that there must be an object identity decision for the phenomenon to make sense is clear.

1.4 Cross-modal information in the real world.

Prism adaptation and the ventriloquism effect show there are laboratory situations where the samples come from different modalities and the object identity decision must be made. Before turning to the next example, note that these cross-modal effects have relevance for more ordinary situations. Suppose you are looking at your pen while writing comments in the margin of this article. Do you know, and how do you know that the visual sample you get (from the pen) and the touch sample you get (from the pen) refer to the same object? Or you are lifting a cup of coffee which is out of sight while reading these words. Do you know and how do you know that the visual sample you get (from the words) and the touch sample you get (from the coffee) refer to different objects? The decisions are critical to keep perceptual systems finely tuned as seen with the example of prism adaptation, and to maintain an immediate experience of a coherent world as seen with the ventriloquism effect.

As in apparent motion, these laboratory curiosities and parlor games provide a window into a routine process that would otherwise proceed unnoticed. The additional conflict between spatial positions created by the demonstrations allows one to see the impact of an object identity decision. Note that conflicts in spatial locations can

also occur naturally. For instance, you have a cold and one of your ears is partially clogged. As you talk to your friend, you have a visual signal for one location (from the friend) but an auditory signal for a different location (from the friend). Do you know and how do you know if the signals refer to the same object (the friend)? Or you are a male adolescent and you pick up your sneakers. The distance between your shoulders has been rapidly increasing and the position sense of your arms has not caught up. Your sense of where things are through touch uses information calibrated for a smaller-sized body, and you localize the sneakers through touch in a different place than you do with vision. Do the visual and touch signals that differ from one another in location come from same sneaker? But whether or not there is a conflict between spatial positions, the object identity decision is necessary. There is already a "conflict" between modalities in the sense that every time more than one modality is used, there is more than one qualitatively different piece of information, and deciding if these different samples refer to the same object or not becomes necessary.

1.5 Stereopsis.

 In apparent motion, the two samples come from two different times, and for prism adaptation and the ventriloquism effect, the samples come from two different modalities. A third type of example may even be when the samples come from different eyes. It is well known that the left eye and the right have slightly different, yet overlapping, views of the world; the distance between corresponding points on the two eyes is the binocular disparity and can be used as a powerful cue to depth. Items close by produce large disparities whereas items far away produce smaller disparities (between the observer and the horopter).

Working backward from the size of the disparities allows distances to be reconstructed. Although the study of depth perception, like the study of cross-modal perception, is typically conducted separately from object perception, a hidden object identity decision is critical for getting any binocular vision.

 I argue the problem of identity stems from what is usually known in this domain as the "correspondence problem". The amount of disparity provides information on depth, but in order to measure disparity you have to know *which* points to measure disparity from. That is, each retina contains numerous points of different lightnesses. One must first identify which point on the left retina goes with which point on the right retina - "what goes with what" - before the disparity between a pair of points can be measured. Consider that "what goes with what" actually derives from the object identity question. If there is a point on the left retina (x) and a point on the right retina (y), what it means to say they should "go together" is that both points were caused by the *same object* in the world. If the two points are caused by two different objects, then those two points should not go together. It can be argued that a same object judgment is required before a sensible disparity between two points can be measured and used to determine distance.

 The importance of object identity may be clearest from the work of Marr (1982), one of the few investigators who emphasizes properties of objects for solving correspondence. Marr identified three constraints at the "computational level" that any solution to the correspondence problem must adhere to which also decrease the computational complexity of the correspondence problem. 1) Black dots can only match black dots 2) one dot from one retina can only match one dot from the other

retina and 3) the disparity must vary smoothly. Marr argues that the rules are derived from the physical situation. For instance, when referring to the first constraint: "...if the two descriptive elements could have arisen from the same physical marking, then they can match. If they could not have, then they cannot be matched" (pg. 114). He argues that the relevant constraints on the physical world are that one point on a physical surface occupies a unique position in space at one time, matter is separated into objects, and surfaces of objects are usually smooth.

That is, all the psychological rules are arguably conditions which increase the probability that dots that will be matched come from the same object. One point on an object cannot be two different colors. Therefore if one member of a potential pair of dots is black and the other white, then the probability that the two samples come from the same object decreases whereas if they are both black then the probability they come from the same object increases. One point on an object can only be in only place at one time, the same constraint essential for cross-modal mappings. Therefore the probability that two points on one image and one point on the other come from the same object decreases, whereas the probability that just one point from each eye come from the same object is higher. As Marr notes, in general the surface of an object varies smoothly in the sense that the texture (depth) variations within an object are small compared to the overall distance between object and observer. Therefore if a measured disparity between a tentative pair of points is dissimilar to the disparity of nearby pairs, then the probability that the tentative pair come from the same source and should be matched decreases, whereas if it is similar, that probability increases.

This example raises the issue of whether the relevant source in the world is one point on the object, or the whole object. The overemphasis on single points may derive from the almost exclusive use of random-dot stereograms (Julesz, 1971) in modern stereopsis research, like older apparent motion displays that used points rather than extended contours. In random-dot stereograms, each image by itself does not lead to any recognizable object, but contains only random black dots on a white background. A meaningful object is obtained only through comparison of the two images, where a specific subset of dots is displaced in one image with respect to the other. While Julesz's ingenious displays were ideal for demonstrating that meaning within single images is not *necessary* for stereopsis, they are less than ideal for investigating how stereopsis *usually* occurs. Random-dot stereograms can take unnaturally long to see any object or depth, sometimes several minutes (Frisby and Clatworthy, 1975). Thus, algorithms based on individual dots may not be exploiting all of the opportunities available to more natural displays. Note that even when applied to individual dots, Marr's computational constraints are still generally derived from the properties of whole objects, e.g. objects vary smoothly relative to viewing distance. Humans evolved rules about identity at the level of "Spelke-objects" - typically midsize bounded solids. Stimuli that are point sources or otherwise degraded can tap into these rules.

In sum, the same object identity question may be present when the samples come from two different eyes, as it is with different times and different modalities. Stereopsis requires a judgment of "same object", like apparent motion, ventriloquism and adaptation. Although laboratory explorations of stereopsis tend to isolate this

system from the other depth cues with which they intermix in everyday perception, the real-world importance of stereopsis for providing information about depth and distance is clear.

1.6 Priming.

Priming refers to the facilitation of the identification of a stimulus (the "target") as a result of prior presentation of a stimulus (the "prime"). For instance, recognizing that a printed letter string is a word of English is faster if the same stimulus is shown briefly a half a second earlier. Various priming paradigms have been used extensively both in language processing to investigate word recognition and the structure of the lexicon (e.g. Forster and Taft, 1994), and in memory to investigate "implicit" processing (Schacter, 1987, 1992). Priming paradigms and manipulations have included presenting primes that are so brief as to remain out of awareness, longer duration primes, ISIs in the iconic storage range, longer ISIs, different fonts of target and prime, different letter case (capital vs. small) of target and prime, target and prime words that differ by one letter, different words in the same semantic category, and the same word in different modalities. Dependent measures for identification have included determining whether a string of letters is a word of English or not, naming the word as quickly as possible, and completing a stem of one or more letters with the first word that comes to mind. With few exceptions, the stimuli are verbal materials. In addition, the studies of language processing and implicit memory rarely make contact with general principles of object perception.

However, according to the current analysis, the formal structure of priming is similar to the other phenomena discussed. Two different samples, in this example separated in time, are processed and should be compared to one another. Is the object identity decision made here too? Although dozens of priming studies have been conducted without consideration of object identity, recently, Kahneman et al. (1992) indeed argue that obtaining the facilitation requires that the target and prime be interpreted as different states of the same object.

The target in their repetition priming experiment was one of 9 letters. Subjects were required to say the name of the target letter as rapidly as possible. To manipulate object identity, the investigators presented two primes in two different locations rather than the typically presented single prime. The target was shown in one of those two locations and could either match the prime letter in the same location, match the prime letter from the other location, or be a different letter entirely. They found that the bulk of the facilitation (27 msec) compared to the different letter condition was only for a prime that came from the same location. Facilitation of the target was minimal (3 msec) when the letter was presented before but in a different location. They interpret these results as suggesting that the benefit of priming is largely due to putting the prime and target in the same "object file", which occurred when prime and target were in the same location. They went on to show that it is not same location *per se* that leads to facilitation, but rather the role same location plays in indicating the two glimpses are the same object. When "same object" is indicated by motion (real or apparent) of a letter, facilitation is obtained even when target and prime are in different locations. The small effects they get for the different object primes - when prime and target are apparently put in different object files - are referred to as the nonspecific preview benefit. Note, however, that these advantages too might be due to object

specific effects where the criteria used to indicate "same object" are "weaker" than those usually accepted, as also occurs in apparent motion (e.g. Shepard, 1984).

For priming, the results suggest that the standard account of the facilitation is not correct. As Kahneman et al. point out, the view that a prime causes node activation is not supported. The data suggest that the facilitation involves backward "reviewing" from target to prime rather than processing in a forward direction from prime to target. Since the decision about whether the two stimuli are the same object cannot be made until after the second stimulus is seen, the view that it is the first stimulus that primes the nodes and causes facilitation cannot be entirely correct. For the present purposes, the striking conclusion is that there is another familiar and highly used laboratory phenomenon, typically studied separately than the others, that requires the same object decision to "work".

Laboratory priming involves object identity. But does it serve any function for everyday perception? Some types of priming occur with long delays of seconds and minutes[2]. While integration across short temporal gaps may simulate what occurs with eye movements, as noted for apparent motion, integration across longer temporal gaps may simulate what occurs with movement of the head and body. A link may be seen with the Gestalt principle of the

Set Effect, one of the original Gestalt grouping principles (Wertheimer, 1923) that has not received as much attention as principles such as Proximity and Similarity. A classroom demonstration is to first show half the observers a picture of a simple line drawing of a man's face, and the other half a line drawing of a mouse. Everyone is then shown an ambiguous picture of a mouse/man and indicates what they perceive. Observers who previously saw the man, see the ambiguous display as a man, whereas observers who previously saw the mouse, see the ambiguous display as a mouse. Having seen the same object moments prior leads to a facilitation in the recognition of that object, as in the priming paradigms. Objects endure. It is a good bet that an object around before is around again. Determining whether a sample now is from the same object as a sample registered moments earlier before turning away, can hasten and make more accurate the extraction of meaningful entities from the current retinal image.

1.7 Gestalt grouping principles-proximity and similarity

The final example suggests that samples from space can be considered analogously to those from different times, modalities, or eyes. The Gestalt principle of Proximity refers to the mental imposition of structure such that items nearby are grouped together, while Similarity describes the structure where similar items are grouped together. Laws of proximity and similarity are often stated as principles, but Gestalt grouping is also a phenomenon. It refers to an experience of spontaneous grouping in a display - that one thing belongs with another. With Gestalt displays, there are also separate samples, such as multiple spatially distributed individual dots and circles as in Wertheimer's (1923) original displays. In

---

[2] Most of the time intervals used in the Kahneman et al. study were within the range over which apparent motion may occur, although facilitation was observed in conditions where an experience of motion did not occur. The longest intervals used were a prime duration of 1 sec. followed by an interstimulus interval (ISI) of 590 msec (study 4). Optimal stimulus and ISI durations for apparent motion are typically shorter than that, but apparent motion with a stimulus onset asynchrony (time from onset of first stimulus to onset of second stimulus) of at least 1 second (see e.g. Farrell and Shepard, 1981) but note that the ISI for Farrell and Shepard was 0.

addition, it appears that what is behind the grouping is a judgment about objects. It seems reasonable to suppose that Proximity works because points on an object tend to be near each other; consequently when working "backwards" from the ambiguous retinal image, it is a good bet that points near one another come from the same object. For Similarity, it may be that the stuff that constitutes an object is more likely to be the same within an object than between objects and consequently similar points on the retina are more likely to come from the same object than dissimilar points. Gestalt grouping[3] may provide another example of the object identity decision. There are distinct samples, A and B, and the question as to whether they refer to the same object is critical. A will be seen as going with B if and only if they are judged to come from the same object.

In Gestalt grouping, the samples are separated exclusively in space, unlike the other phenomena discussed thus far. A potential problem is that in vision, spatial separation is seen as definitive with respect to specifying multiple objects (see Kubovy, 1983, 1988). Kubovy discusses the point with the example of Tweddledee and Tweddledum; they are alike in all ways, yet we know they are different people because they occupy different locations at the same time. Kubovy argues that two colors in the same location specifies one object, but one color in two locations indicates two objects - space is an "indispensable attribute" for

numerosity[4]. Based on the theory of indispensable attributes, it would appear that an object identity decision is not needed when there are discrete visual samples separated exclusively in space. However, consider the prismatic displacement situation where the arm is made to appear in a different location from where it feels. The differing spatial locations at the same time can nonetheless refer to one and the same object. Even though space is an indispensable attribute for at least one modality, if the two samples come from the medium of space at the same time, a decision as to whether they reflect the same object or not is still needed. Entirely within vision, one-object interpretations of two discrete simultaneous spatial locations include seeing a reflection of an object through a mirror, and experiencing "double vision" (diplopia).

Related to this issue is that the two samples in Gestalt grouping appear to come from different parts of the same object, rather than (the same part of) the same object. This solution can occur in other domains as well. If vision and touch signals are in different locations, they can specify one very fat arm that extends through both locations; the touch signal comes from one part, and the visual signal from the other (suggesting a fourth possible interpretation to the spatial discrepancy between modalities discussed earlier.) The question of whether two samples come from the same object and the question of whether two samples come from different parts of the same object may branch off into different decisions, but they begin with the same input.

---

[3] The other Gestalt principles (good continuation, closure, common fate, Pragnanz, set effect) may be understandable in similar ways,. e.g. discrete entities which do the same thing at the same time (common fate) may be more likely to come from the same object than those which are not, as with synchrony (section 1.3; see also Radeau and Bertelson, 1974; Radeau, 1994).

[4] Kubovy's theory of indespensible attributes includes comparison to audition where it is pitch, not space, which is indispensable. One pitch in two locations is one auditory object, but two pitches in one location still retain their separate identity.

1.8 Discussion.

The six examples discussed show that an object identity decision is required in diverse well-known situations when samples come from different times (apparent motion, priming), different modalities (prism adaptation, ventriloquism effect), different eyes (stereopsis), and different spatial locations (Gestalt, apparent motion), whether under the name of unity assumption, pairing, correspondence, object files, or identity. Thus, object identity is a required computation not just when samples come from different times, but also when they originate from other sources. Moreover, identity applies to samples that can vary from small to large[5]. A spatially isolated small "point" can be a sample, such as in a simple apparent motion display, as can a contour extended in space, i.e. a form. Similarly for time, a sample can be an isolated "point" in time, as with a stimulus that is present only for an instant, or have an extended contour in time, i.e. an event. Finally, the laboratory phenomena all reflect the types of processing that occurs in everyday perception and apprehension of the world; the object identity decision is more generally critical for tracking objects over time, separating objects from one another, keeping perceptual systems accurate, getting quantitative depth information, and experiencing a coherent world, accomplishments usually considered separately. The abundance of samples requires the same decision be made in many different domains and made frequently.

Will the criteria of object identity always be the same? The different domains are at least partially modular. Different brain regions and psychological "isolation" evidence (Weiss, 1941) suggests that each system can function independently. This could be taken to mean that the different problems are only coincidentally similar, or different content areas may suggest that criteria appropriate for one area are useless for another. Yet parsimony suggests a solution common to all be considered. A good solution tends to recur in different problems with the same formal structure (e.g. Dennettt, 1996, Shepard, 2001), for example, the high frequency of hexagons in nature. If one abstracts away from the specific content, the object identity problem is the same in different areas. Unrelated problems can have the same solution because of independent evolution (e.g. vertebrate eyes, Dawkins, 1996) or "blueprinting" (Rozin, 1976) where a solution evolved for one problem then became accessed by now independent brain regions. Alternatively, there may be less modularity of the domains than typically thought (see Bedford, 1997, and Shallice, 1988 for problems with double-dissociation methodology).

A final caveat concerns consciousness. An object identity *decision*, does not here imply that conscious deliberation on whether two samples do or do not refer to the same object occurs. Conscious awareness is not necessary. Whether awareness can nonetheless be influential is a different question. Conscious beliefs that one is in the presence of one or two objects appear to sway the "decision" in several domains discussed. It would be intriguing if conscious penetrability were possible in all manifestations of object identity, and identity even the gateway through which

---

[5]Avoiding circularity in the selection of samples (how does one extract two different samples from the flux, and then turn around and ask whether they really are caused by the same thing?) can be done by using discontinuities for continuous dimensions like space or time.

consciousness can influence perception, but I stop short of making such a claim.

Part II: A solution common to all modules

What properties would a general solution for object identity need? 1) There is no single necessary and sufficient property that solves the identity problem. For instance, if you look at your right arm, the fact that both touch and sight signals come from the same location at the same time seems to indicate that the signals come from the same arm (Held, personal communication, see "Held's paradox", Bedford, 1994). But if the visual signal is displaced with a prism, the different locations of the visual and proprioceptive signals *still* lead to the same-object decision or prism adaptation would never occur. When properties such as location differ for two samples, the samples will sometimes be judged to refer to the same object, and sometimes not. The solution must be flexible. 2) The solution must be abstract enough to apply to time, space, and modality as well as numerous different content areas.

The solution is an entire hierarchy of geometric criteria. If there are two samples, the more overlap, the greater the probability that they will be judged to refer to the same object. But if two samples are not identical, how can one determine how different they are? At issue is a metric of what it means to be "similar", a thorn in the side of many psychological problems (see e.g. Medin, Goldstone and Gentner, 1993). I suggest here a hierarchy than can serve as a metric of similarity specifically for the object identity problem. Most of the samples in real-world situations have extended contours - that is, "forms". When the samples are minimal in size (individual points), they can be considered limiting cases - degraded forms (see also Appendix 2).   Two forms that are

not identical have been compared for similarity in the psychological literature by the intuitive features they share (e.g. number of corners), by spatial frequency components, and with ecological optics, to name a few strategies. However, the most natural candidate is geometry, the abstract study of form. In the present theory, the criteria of object identity in all domains come not just from one geometry, but from an entire hierarchy of geometries that was formulated by mathematician Felix Klein one hundred years ago.

What follows is the general framework of the solution, an informal description of the geometries in the hierarchy (see also Appendix 1) along with what they represent in cognitive/ psychological terms, application to the varied domains from Part I, and discussion.

2.1 General framework of the solution - transformations

The transformation approach to geometry is particularly relevant for object identity. The traditional approach is based on *axioms*, in which all properties of the geometry are derived from a small set of assumptions. Euclid proposed 5 postulates, 5 "common notions", and 23 definitions which lead to all the theorems and properties of the geometry that bears his name, the one and only geometry for two thousand years beginning in 300 B.C.. Klein (1893/1957) originated the *transformation* approach to geometry. He showed that geometry could also be characterized by transformations, such that the properties of the geometry are those properties which remain *unchanged by a group of transformations*. He further showed that different groups of transformations lead to different geometries in addition to Euclidean geometry. The resulting geometries could be ordered based on the number of properties left unchanged

by the defining group of transformations: the more radical a transformation, the more properties that are altered, the fewer the properties that remain in the geometry, and the larger/more general the resulting geometry.

To apply the hierarchy to object identity, the two samples will be considered as two forms which differ from each other by a transformation (which corresponds to a geometry). For each point on form 1, one can identify the corresponding point on form 2 to which it gets mapped by the transformation. The more properties that are altered by the transformation, the less likely those two samples will be judged to refer to one object. Intuitively, the more geometric properties that are altered, the less "similar" the two forms become. The strongest criteria, whereby two samples are judged very likely to be one object, corresponds to the smallest and richest geometry within which few distinct forms are considered identical. As the criteria get weaker, and samples are judged increasing less likely to be one object, they correspond to increasingly larger geometries within which more and more distinct forms are considered identical. How this general ordering leads to a specific decision about two samples will be seen later. The geometries all nest within one another such that each larger one is a superset of all the preceding ones. Figure 1 shows a schematic of the nesting along with properties within each geometry. The particular geometries, in order from smallest to largest, are Euclidean, Similarity, Affine, Projective, and Topological geometry. Each contains properties that correspond to meaningful features in psychological terms.

2.2 Core of the solution - five Geometries.

2.2.1 Euclidean Geometry - ignore location. The most familiar geometry,

Euclidean, is generated by the group of transformations known as isometric transformations; these can alter the location of a form in its entirety, as if the form were simply picked up and dropped somewhere else. For instance, a square could be moved a few inches sideways, or rotated (see section 3.5 for more on rotation). The only property altered by an isometric transformation is absolute position. All other properties of the square remain unchanged, such as the distance between any two points on the square (i.e. length) before and after the transformation, the angles of the square, the parallelism of the two pairs of sides. All the properties which remain unchanged by isometric transformations are in Euclidean geometry but those that are altered are not. *Thus, absolute location is not a property in Euclidean geometry.* If a square is taken and "transformed" (which we call "moved") to a location 3 inches away, the square and the resulting form (which we also call a square) are equivalent forms in Euclidean geometry. Because absolute location is not in the geometry, it is irrelevant for distinguishing between forms. But if a square is put in a vice such that the top part is compressed and the sides are no longer parallel (for example), then the square and the post-transformation form are not equivalent forms in this geometry; parallelism *is* a property in Euclidean geometry.

Because it allows the fewest possible changes to a form and contains the most properties, Euclidean geometry constitutes the strongest criteria of object identity. Two samples related by an isometric transformation will have the highest probability to be judged as coming from the same object. Note that Euclidean geometry largely corresponds to our intuitions, where objectness "feels" independent of position. It is typically believed to describe the natural

world in which we live at the scale at which we operate, and our language seems to reflect this. As noted above, a square and a new form that is 3 inches away but otherwise identical is also called the same thing, a square, even though strictly speaking, and importantly, it is *not* identical: its absolute location is different. A square that is only transformed by changing its absolute location gets a special name: it was "moved". "Transformation" in our language is reserved for transformations that alter those properties that we intuitively think of as changing its "squareness", namely those properties in Euclidean geometry, which are not altered by isometric transformations.

2.2.2 Similarity geometry - ignore size. However, Euclidean geometry is only a special case of the more general Similarity geometry. Similarity transformations, which generate Similarity geometry, allow uniform expansions and contractions to both dimensions of a form provided they are applied equally to both dimensions, e.g. magnify both horizontal and vertical directions three-fold (x'=3x, y'=3y). Unlike isometric transformations, similarity transformations do not preserve the distances between pairs of points (lengths), but they do preserve the ratio between pairs of lengths. Thus, they are more radical and change more of the original form than isometric transformations, yet retain many properties of the original form as well. Figure 2 shows a few sample forms before and after similarity transformations, as well as those that generate the other geometries.

Similarity transformations allow any changes which change distances with uniform ratio. What remains after the transformation is a scale model of the original. That is, similarity transformations allow the *size* as well as location to be changed. In Similarity geometry, not only are a square and a displaced square

equivalent, but so are a square and a bigger/smaller square. Size and location are altered by similarity transformations, and therefore these properties are not in Similarity geometry. Whereas Euclidean geometry excludes the property of location, Similarity geometry excludes size as well as location. Whereas size is a property of shape within Euclidean geometry (since it is preserved by isometric transformations) size is not a property of shape within Similarity geometry (where it stands outside along with position).

It is clear that transformations which can alter both size and location are more radical than transformations which alter only location, but it is important to note that the relation between these two levels is not arbitrary. One can change the absolute location of a point without changing the size of the form, but it is not possible to change size without also changing locations of individual points. That is, changing lengths between pairs of points changes positions of individual points, but positions of individual points can be changed without changing the length between pairs of points. Size presupposes location, but not vice-versa. Each more radical group of transformations as the levels ascend alter additional properties that cannot be altered without other properties that they presuppose.

While for the most part intuitions mesh with Euclidean geometry, they also appear to extend to Similarity geometry. A square and a small square are still both called *squares*; our intuition is that the "squareness" is still there. Size is sometimes considered a property of shape, and at other times not (see section 3.4 for more on this duality). In the present theory, two forms that differ by a similarity transformation will be judged less likely to originate from the same object than two forms that differ by an isometric transformation. Yet two forms which differ

in size will in many circumstances be judged identical, as they are in Similarity geometry.

2.2.3 Affine geometry - ignore angle.

Within the third more general level of Affine Geometry, an even greater variety of forms are identical. Affine transformations also allow uniform expansions and contractions of each dimension, but now those changes do not have to be the same for both orthogonal dimensions. For instance, a square can be turned into a rectangle by magnifying the horizontal direction (x'=3x) while leaving the vertical direction unaltered (y'=y), or changing it by a different amount, e.g. (y'=2y), or (y'=.5x). The square can also be pulled uniformly from one diagonal (i.e. in rectangular coordinates x depends on *y*, e.g. x'=3y) to produce a rhombus. Note that Similarity geometry was a special case of Affine geometry where both dimensions are altered equally. In the more general case, allowing uniform changes separately to each dimension typically alters the *angle* of the form. In Affine geometry, angle is added to the list of properties that are not contained within its study. Note that the relationship of angle to size is again one of presupposition, or containment. The size of the form can be changed without changing the angles, but the angles of a form cannot be changed without changing its size. In Affine geometry, a square, a rhombus, a small square, and a square located 3 inches away are one and the same form. So too in object identity, I argue two samples that differ by an Affine transformation can be judged to refer to the same object. They will be less likely to be from the same object than the less radical transformations which alter fewer properties, but still more likely than the 4th level, which alters even more properties that give a form its distinctness.

2.2.4 Projective geometry - ignore parallel lines.

The fourth level results from projective transformations which can destroy another property: parallelism. With projective transformations, lines which are initially parallel can become non-parallel and vice versa. For instance, a square can be turned into a trapezoid with appropriate tugging. Within Projective geometry, now even a square and a trapezoid are indistinguishable, along with the other quadrilaterals described within the smaller geometries. Parallelism presupposes the properties from prior levels. Changes in angle can occur without changing the non-parallelism of the two line segments, but changing the parallelism of two lines destroys what is meant by angle. What properties are left in the geometry? Collinearity remains unaltered; if three points before the transformation are in a straight line, they will be after as well. Order of points is also unaltered. Thus, some properties of the initial form are still recognizable following projective transformations, though we are stretching our intuitions to see the similarities.

2.2.5 Topology - ignore straightness.

Intuitions are stretched to the limit, if not beyond, with Topological geometry. Here, the transformations are radical enough to allow squares to be transformed into circles. The "straightness" of the sides of a square are not preserved because topological transformations further allow collinearity to be destroyed. Topological transformations permit nearly unlimited pushing, pulling, stretching, bending and deforming to a form, provided that points which are continuous are not split and made discontinuous and points initially distinct not glued together. Thus the property of being a closed curve is one of the few things that remain from the original form. Within Topological geometry, forms we would never consider the same at an intuitive level are identical. A square and a circle are the essence of what we mean by different forms, one has "squareness", the

other "circleness". Yet within Topology they are the same, and for object identity they arguably can be as well (see also Chen, 1985 and section 3.1). Two samples which are related topologically will be much less likely to be judged to refer to the same object than the other smaller stronger levels, but still more so than if topological properties are destroyed.

2.2.6 Beyond topology- ignore connectedness. Non-topological changes are so radical as to be officially beyond the hierarchy of geometries. Topological geometry is the last level in Klein's hierarchy. If any other properties are destroyed, nothing that geometers consider to be shape is preserved, even under the most liberal definition of shape. Non-topological transformations would allow many to one mappings; i.e. fusions where distinct points are made non-distinct, and one-to-many mappings, where points are split and continuity turned into discontinuity. Samples that differ from one another in such ways would be very unlikely indeed to be judged to originate from the same object.

2.3 Applying the theory - space, time, modalities.

To summarize, in the present theory, the object identity decision is probabilistic such that the more radical a transformation, the greater the properties of the original form are altered, and the less likely the pre-transformed and post-transformed samples will be judged to refer to the same object. In general, two samples which differ by an isometric transformation are more likely to refer to the same object than if they differ by a similarity transformation, which in turn is more likely that if they differ by an affine transformation. Affine changes are more likely to reflect a single object than projective changes, which are more likely than topological changes, which finally are more likely than non-topological transformations. Observers possess the geometries in the sense that if it is decided that two samples physically distinct are nonetheless the same (object), then one is within and governed by the geometry in which those two forms are considered equivalent. While intuition is largely constrained by Euclidean geometry, core processing of identity is not likewise limited. In practice, this general probabilistic information can be used for an actual decision - Yes or No - *by choosing the mate for a sample which is related by the lowest level transformation available in the situation*. Thus the same pair of samples can be judged to refer to the same object in one situation, yet judged to be different objects in a different situation.

To see how the theory can be applied, each domain of object identity from Part I will be considered. In apparent motion, there are two forms which differ by a geometric transformation. The findings about rigid and plastic transformations in the literature now become sensible. As noted earlier, researchers have found that rigid motions occur frequently between the two views. For instance, a square and a square a few inches away are resolved by interpolating a lateral shift of the whole square rather than a deformation which reforms back into a square. A form that is a complex polygon followed by another will settle into a rigid rotation of the initial form to a new orientation when possible, rather than a polygon which deforms into another. (Shepard and Judd, 1976; Shepard, 1984). These perceptions seem sensible and "simplest". Simple for object identity can now be understood more rigorously as level 1 of the hierarchy. Rigid motions are precisely isometric transformations in which only the property of absolute location is

altered. The theory predicts they should be preferred.

Predictions are more rigorous in the "competing motion" paradigm (Ullman, 1979), where a sample is presented at time 1, but two samples, a and b, are presented at time 2. The dependent measure involves assessing whether the original sample will move to stimulus a or stimulus b. That is, when there is a choice of which stimulus will be judged to refer to the same object as the glimpse at time 1. In the present theory, the two samples which differ by an isometric transformation will be judged more likely to refer to the same object than the two samples which differ by more radical transformations. When there is a choice between a higher and lower level transformation, apparent motion will occur between the lower-level pair. This is precisely what occurs, usually described as an effect of shape. Mack et al. (1989) found that a stimulus "Z" is seen as moving to an identical "Z" in another location rather than to a "E" in an equally distant location. The stimulus maps to the "same shape" rather than to a "different shape" which can be redescribed as a preference for an isometric transformation (level 1) rather than the much more radical non-topological transformation ("level 6") which relates Z and E.

Moreover, there is evidence that when an isometric option is unavailable, the lowest level transformation that is available governs apparent motion. Chen (1985) pitted a topological transformation (the present level 5) against a non-topological transformation ("level 6"), and found apparent motion between the topological pair. For instance, in one of Chen's displays the initial stimulus was a solid circle located in the center and the two stimuli at time 2 were a solid square to the left of center and a solid circle with a hole in it to the right. Although a circle and a circle with a hole in

it seem to some observers intuitively more "similar" to each other than a circle and a square, motion was seen between the circle and the square. A circle and a circle with a hole in it are not topologically related to each other, but a circle and a square are. This is precisely what is predicted from the other end of the hierarchy - two samples which differ by a level 5 transformation should be judged more likely to result from the same object than two samples which differ by a transformation that is too radical to be considered within geometry ("level 6").

As Chen and others (e.g. Kolers and Pomerantz, 1971; Shepard, 1984) have noted, and as the experiments described above suggest, non-rigid plastic deformations are also experienced in apparent motion even though rigid motions (isometric) seem preferred. If the two glimpses are a square and a circle, the square will appear to contort and transform into the circle. Shepard argues that when rigid resolutions are not possible, plastic deformations are seen because the system continues to identify the two different views as glimpses of the same object by using "lesser criteria of object identity". The geometric hierarchy provides a rigorous way to specify what those lesser criteria are. While the existence of both rigid and plastic experiences may appear contradictory, they are not if multiple levels of a hierarchy are accessed for object identity, rather than a single necessary and sufficient criterion.

What about the contradictions in the literature concerning shape? As noted earlier, some investigators report shape is influential in apparent motion, but others argue it has little or no effect. While appearing contradictory, the investigators who have drawn those conclusions are really asking different questions. One question is: Can changing the shape of the stimulus be

shown to have an effect on apparent motion, to which the answer is "yes" ("shape matters"). The second question is: Can apparent motion occur even when the shapes are made different, to which the answer is also "yes" ("shape doesn't matter"). They are both correct. Differing "shape" samples can still lead to apparent motion because of the many levels of criteria of object identity. Samples which differ by the non-isometric transformations of levels 2-5 (similarity and so on) can still be judged to refer to the same object because of the geometric properties they still share; one object will still be seen. Yet closer inspection reveals that the transformations with more geometric properties preserved are more likely to be judged to refer to the same object than those with fewer geometric properties preserved. Such an effect of shape can be revealed through choice of motion direction and/or experienced goodness of motion between two stimuli (see Dawson, 1989).

As discussed, there is evidence from existing experiments that isometric geometry is preferred to other levels ("same shape, different shape") and evidence at the other end of the hierarchy that topological geometry is preferred to non-topological changes. The theory also predicts that intermediate levels should be discernable. That is, motion between similarity transformations (level 2) should be chosen over motion between affine transformations (level 3), affine transformation should be preferred over projective ones, etc. Using a variant on the competing motion technique we obtain generally confirming evidence (Bedford and Mansson, under review). The only exception concerns level 2, similarity transformations, which seem less preferred for some of the subjects than predicted by the hierarchy. However, all the similarity transformations used map the original form to a smaller form and therefore to a stimulus

further away than the other choices. When distance between nearest edges is controlled for (see section 3.2), level 2 falls back into the proper place.

The second example of object identity argued for in part I was prism adaptation (section 1.2), where the two samples come from two different modalities (vision and proprioception) at the same time rather than two different times within the same modality. The theory can be used here too to determine if the two samples derive from the same object. To do so, consider space as a large form. A geometric transformation relates the visual form (or space) to the proprioceptive form (or space) For instance, in a standard prism adaptation experiment, the primary distortion of a 20 diopter prism with the base facing left is a rigid 11.3 degree leftward shift of the visual image. That is, a wedge prism creates predominately an isometric transformation between visual space and proprioceptive space. If an observer looks at her arm through the prism, only the absolute location of the arm is altered.   All other geometric properties of the arm remain unchanged, size, angles, parallelism, order of points, connectivity etc. In an isometric transformation, all the points are shifted by the same amount to "produce" proprioceptive space; because only absolute location is different, the proprioceptive and visual samples are very likely judged to refer to the same object, and the error correction known as prism adaptation results.

If the two samples differ by a transformation more radical than an isometric transformation, according to the theory they will be less likely to be judged to refer to one object. Since adaptation is critically dependent on the decision that the two samples are the same object, interfering with this decision should interfere with the amount of adaptation. Ascending the

hierarchy should lead to less and less adaptation.

While the majority of systematic studies of mappings between perceptual modalities have focused on uniform shifts (isometric), Bedford directly compared five levels of transformations (1993a, 1994). Using a paradigm developed by Cunningham (1984, 1989), the visual positions are specified on a computer monitor and motor positions are indicated by position of a pen on a digitizing tablet. When a pen is moved around on the tablet, a cursor representing the pen is shown visually on the screen, like with a computer mouse. Distortions between visual space and motor space are created with software rather than with prisms or lenses. For instance, a mirror transformation could be obtained by having the visual cursor move to the right whenever the pen was moved to the left. For the isometric transformation, all of motor space was shifted uniformly down with respect to visual space. The similarity transformation shrunk both x and y dimensions of motor space equally which effectively turned a visual square into a smaller square in motor space; small motions on the tablet were uniformly expanded on the screen. The affine transformation turned a visual square into a sheared square in motor space, and the projective transformation turned a visual square into a motor trapezoid. Finally, for the topological transformation, straight lines were made curved and a visual square space turned into a small circular motor space. The data suggest that the first four levels order themselves from easiest to hardest to learn. The fifth level was unexpectedly easy to learn, but closer inspection revealed that subjects learn the scale change (level 2) and not to the circular components (level 5).

Bedford (1993a) also used atypical mappings directly with a prism adaptation task rather than the more abstract computer mapping task. A prism adaptation paradigm typically uses a single dimension of space from left to right, rather than two dimensions. While two dimensions of space define a form, one dimension of space defines a line. A line can either be moved over rigidly (level 1), can be uniformly expanded or contracted (level 2, 3, and 4), or non-uniformly stretched with different parts of the line expanded more than others (level 5). For one dimension of space, the second, third, and fourth level of the hierarchy, similarity, affine, and projective collapse to one level as shown in Figure 3. Non-topological mappings ("level 6") would split a line into two lines, or change the order of points along the line, or fuse distinct points. The experiments found 1) a (non-linear) topological mapping produces more adaptation than a many-to-one non-topological mapping (Bedford 1993a) 2) a preference for level 1 rigid shifts over all other changes, as indicated by the pattern of generalization following limited input (1989) and 3) a preference for level 2,3,4 uniform stretches over more radical changes when rigid shifts aren't possible (Bedford, 1989, Bedford 1993b). Older studies on prism adaptation have found that the small non-rigid distortions produced by a prism are less well accommodated than the rigid ones (See Welch, 1978). Thus with prism adaptation, like with apparent motion, a geometric approach to object identity provides a useful way to understand, organize, and predict.

When there is a conflict between vision and audition, as there is in the third example of the ventriloquism effect (section 1.3), time is an additional consideration. Auditory stimuli unfold over time - the long whistle of the tea kettle, human speech, the pulsing of white noise in a minimalist ventriloquism experiment. Geometry is the study of form, which has a natural application to space.

Yet geometry can also be applied to time. One can think of extensions in time analogously to extensions in space, temporal contours instead of spatial contours. Samples extended in space comprise forms and objects; samples extended in time constitute events. To apply the geometric hierarchy to time, one can use the physicists' trick of transforming a problem to an easier domain, working it out there, and transforming it back. Time can be converted into space where extensions in time become extensions in space; a point *later* in time maps onto a point to the *right* in space. Following this conversion, what remains is a line, and the geometric hierarchy is reduced to three levels as shown in Figure 3. That is, the levels are identical to the levels described for one dimension of space discussed in the previous section. The geometry of a one-dimensional line is less complex than that of a two dimensional form. Any mysterious properties of time are absent when viewed with the geometry of a line in space.

As discussed, the least radical group of transformations for a line shifts the entire line over to the left or right (isometric). The second level allows uniform expansions and contractions, which can make the line uniformly larger or smaller (similarity, affine, projective). The third allows non-uniform expansions and contractions (topological). If the line is broken and turned into two lines, or the order of points are intermixed, or distinct line segments become one, then a non-topological transformation has taken place. To convert back from space into time, moving the entire line to the right or left corresponds to moving the entire temporal episode later or earlier in time. For instance, if I have appointments with Tom, Merrill, Paul, and Bill back to back for 1 hour each, beginning at 9:00 am tomorrow, an isometric transformation allows me to

start the visit at 9:40 and push everyone forward by the same amount because my train was late. The similarity transformation can shorten the day - I need to be done by 11:40 to give a Cognitive Science Brownbag lunch talk and all appointments are shortened from 1 hour to 1/2 hour each. The topological transformation allows me to spend an hour with Tom but 20 minutes with Merrill, and the non-topological transformation allows any schedule change including combining the meetings of Bill and Paul, having two different meetings with Merrill, or meeting with Tom in between other appointments.

The theory consequently predicts that the ventriloquism effect should be easier to get if the auditory stimuli are delayed a fixed amount with respect to the visual stimuli than if the duration of the auditory signals are all reduced by the same scale factor with respect to the duration of the visual stimuli. That in turn should be easier than if the durations of some of the auditory stimuli are reduced, but others not, which in turn should be easier than the most radical non-geometric transformations where individual sounds can be turned into two different lights, or two sounds into one light. A same-object conclusion should be increasingly less likely, which should lead to less of a ventriloquism effect.

The timing of the sounds with respect to lights is in fact critical for the phenomenon (Jack and Thurlow, 1973; Radeau and Bertelson, 1977, 1987; Thomas, 1941), as noted in Part I. The "synchrony" of lights and sounds influences whether the visual and auditory signals will be judged to refer to the same object and therefore whether or not the ventriloquism effect will occur. If the lights and sounds are in synchrony with one another, the effect is optimized, but if they are asynchronous, the effect is reduced. In the present theory, several distinct types

of asynchrony would be recognized. One type, in which the auditory stimulus is on periodically, but the visual stimulus is continuous (Thomas, 1941; Radeau and Bertelson, 1987), can be redescribed as a non-topological transformation in which multiple distinct auditory signals are being turned into a single visual stimulus. This destroys any recognizable connection between the lights and the sounds - they become intuitively speaking "unrelated" or "asynchronous". This manipulation appears to completely eliminate ventriloquism. In other conditions, one signal train is delayed with respect to the other (Choe et al., 1975; Radeau and Bertelson, 1977). According to the present theory, this condition is an isometric transformation. As predicted, ventriloquism is easier to get here than with non-topological mappings. The effect is reduced, but not eliminated. Direct comparisons between the different types of asynchrony generated by geometry have not to my knowledge been tested and provide an interesting set of manipulations for exploration. The present theory provides a rigorous way to structure the intuitions behind temporal manipulations as well as spatial manipulations.

For samples that originate from different eyes, as in stereopsis (Section 1.5), the individual points on each retina may not seem amenable to a very useful form analysis. However, with the exception of random-dot stereograms, there are forms on each eye and they appear to be used in the matching process. Without the forms, matching is slow; random-dot stereograms require an unnaturally long viewing time to see depth, sometimes requiring several minutes of viewing with concentration (Frisby and Clatworthy, 1975). In addition, sets of random dots can be analyzed for more global geometric properties. Finally, for individual points, Marr's (1982)

"uniqueness" matching principle described earlier ("Almost always, a black dot from one image can match no more than one black dot from the other image", pp. 115) can be considered to be a comparison between topological and non-topological mappings. The former requires a one-to-one mapping, but the later does not. Mapping one point from one eye to two points from another eye would violate topology, and would be less desirable than other alternatives. Stereopsis seems a promising domain for the same geometric analysis that are applicable to the other domains of object identity.

Even fewer predictions have been tested in priming (Section 1.6). One direction to pursue in letter priming would involve comparing the geometric properties of different letters. Asimov (1995) analyzed that in sans serif font, the letters "C, G, I, J,L, M,N,S,U,V,W and Z" are all topologically equivalent to each other, as are "E, F, T, Y". We would expect then that letter pairs I and J and letter pairs E and T to lead to more priming than pairs I and E or J and T. Another direction would be to use picture priming with simple closed shapes. A circle should prime a square more than a broken up square would, but not as well as a trapezoid should, and likewise for the remaining levels.

For the Gestalt grouping principles (Section 1.7), Proximity reflects variations *within* isometric transformations. All of the comparisons discussed have been *between* the different levels. Quantitative variations within the same level are also easily ordered. Two samples related by big changes in position are less likely to refer to one object than two samples related by a smaller change in position, assuming everything else is held constant (see section 3.2).   This describes the principle of Proximity, where a sample gets grouped with whichever sample

is closest in the situation. It is not surprising that grouping by Proximity also gets applied to a number of diverse phenomena such as the ventriloquism effect (Radeau, 1994), apparent motion (e.g. Green and Odom, 1986), and Pavlovian conditioning (Rescorla, 1985) . These phenomena all involve the object identity decision, and proximity can be derived from the general solution to object identity - the hierarchy of geometries. Note that quantitative variations within other levels can occur as well, although there is no unique gestalt label to describe them.

Within the present theory, the Gestalt principle of similarity is a catch-all category which includes a number of different comparisons both within each of levels 2 through 5 and across all the different levels. Across the different levels, stimuli that are related by a lower level will be more likely to be grouped together than stimuli related by a higher level.

The final application discussed concerns perceptual constancies (see Introduction). Isometric transformations correspond to *Position constancy*, where the position of an object appears constant despite changes in retinal position.. If you move your eyes to the right, the position of the image of the object on the retina change, but nothing else. *Size constancy* corresponds to a similarity transformation. If you walk up to an object, the size of the image on the retina changes, but nothing more radical. The general label "shape constancy", where "shape" is perceived as unchanging despite retinal variations, should subdivide into two categories. *Shape constancy 1* under affine transformations result from looking at an object while you are slanted in depth and very far away. If you look at a rectangle from an "infinite" distance, and turn clockwise in depth, the image will be a sheared rectangle. Practically speaking, this

will result at a distance of approximately 20 feet. *Shape constancy 2* corresponds more generally to looking while slanted in depth at any distance, which causes a general projective transformation. A projective transformation that is not also an affine transformation, e.g. a rectangle that becomes a trapezoid on the retina, will occur at closer distances.

In addition to categorizing the constancies, the theory predicts those from the lower levels are easier to achieve. The different constancies are also not entirely separate accomplishments of separate size, position and shape "systems" which belong in separate chapters of a perception text, but instead share a mechanism that extracts which geometric properties have stayed the same and which have changed. To my knowledge, experiments that compare the ease of constancies have not been conducted, although their connectedness was implied early (Ittelson, 1951). The constancies may be another example of a phenomenon that requires the object identity decision. Like the other examples, achieving constancy seems to require that the varying retinal images refer to the same object; e.g., size constancy refers to the perception that *an object* appears the same size despite changes in retinal image size. But are there different samples? Discrete samples could come from very fast rates of change where intermediate images are not useful. Alternatively, perhaps the very continuity of many constancy situations, with the continuously varying size, angle and other geometric properties has provided the basis for using these criteria in less clear situations.

Discussion
3.1 Why these geometries? - evolution and empirical support

Geometry is the study of form and is therefore relevant. But there are an infinite number of geometries, most with its own group of transformations. Mathematical definition "...is restricted by no other rule then that of avoiding contradiction" (Cassirer, 1944, pp. 4). Euclidean geometry is believed to describe the world at the scale in which we operate. But other geometries described here seem distinctly false; e.g., we do not live in a topological space where squares and circles are equivalent. If the other geometries do not apply to the objects and space that we interact with, why consider them as candidates for object identity? Does it make evolutionary sense to possess rules that do not apply to our immediate world? And why consider these geometries over others that are also believed false of our world such as hyperbolic geometry, where the angles of a triangle sum to more than 180 deg., spherical geometry where the shortest distance between two points is circular, and geometry based on an infinite number of dimensions?

While the geometries in the present hierarchy are false in a certain practical sense, they are nonetheless sensible from an evolutionary perspective in a way that any arbitrary geometry is not. To illustrate, consider again the axiomatic approach to geometry within which a small set of given assumptions form the foundation of each geometry. All properties and theorems are obtained by derivation from those assumptions. In Euclidean geometry, there were originally 5 "common notions" and 5 postulates, such as the 4th postulate which states that right angles are all equal to each other, and the famous 5th parallel postulate: Given a line m, and a point n not on the line, there exists one and only one line that contains n and is parallel to m. All of those axioms are believed true of our world on the scale at which we operate. Some geometries

have axioms which are not. For instance, hyperbolic geometry substitutes the parallel postulate with the assumption that there are at least two parallel lines.[6] But Affine, Similarity, Projective, and Topological geometries do not have axioms which are *false* of our world. Instead, they are *missing* axioms. For instance, in Affine geometry the 4th postulate on angles is not applicable (angles are not invariant in Affine geometry) nor is the 3rd postulate, which is on the circle. Projective geometry additionally does not have any of the assumptions that imply parallel lines cannot cross. As we ascend the hierarchy, each geometry is obtained by removal of more and more of Euclid's assumptions. Similarity geometry contains many of the axioms of Euclidean geometry, but not all of them, and critically, it does not add any new axioms which are false of the world. The axioms are a subset of Euclidean axioms. Axioms that give rise to Affine geometry are a subset of those of Similarity geometry. Projective geometry has even fewer axioms and is a smaller subset, and Topology has the fewest assumptions. There is a difference between axioms which are false and axioms which are omitted. Cheng (1984; Cheng and Gallistel, 1984) argues that natural selection should not lead to the internalization of any geometric property in a systematically wrong way, but could fail to capture or use all the potentially useful rules about the way the world works. Consequently, not living in a topological space (for instance) would not rule out possessing the rules of Topology. Candidates for weaker criteria and weaker geometry that are internalized and relevant for cognition and perception can be

---

[6]While the geometry further out in space may be either hyperbolic, Euclidean, or elliptical, most people believe that Euclidean does provide a very good fit to the local world that we inhabit, despite the controversies that the parallel postulate has caused.

geometries that are subsets of axioms that are believed true of our world[7].

Note that *individual* levels of the hierarchy have been used successfully in perception, and provide some empirical support. Projective geometry is the geometry that characterizes projection of the three-dimensional world onto the two-dimensional retina and consequently has been applied to shape constancy and picture perception (e.g. Cutting, 1987; Niall and MacNamara, 1989, 1990; Perkins, 1974) and extended to color perception (Bergstroem, 1977), impressions of "good form" (Perkins, 1976), perception of rotational motion (Johansson, 1974) and the Ames window illusion (Olson, 1974). Affine geometry has also been applied to picture perception (e.g. Cutting, 1988), and since affine geometry is the most general linear transformation, it has been applied to diverse problems in and out of perception such as perception of structure from motion (e.g. Koenderink and van Dorn, 1990). The extraction of topological properties has been argued extensively by Lin Chen to be a fundamental characteristic of the visual system (Chen, 1980; 1982a; 1982b; 1982c; 1985) including visual grouping, card sorting, texture discrimination, and the object superiority effect. Euclidean geometry is involved whenever exact

---

[7] Note also that nothing that has been suggested is incompatable with what we know of the brain, nor even difficult to implement. For instance, neurons from inferior temporal cortex (IT) have properties that may be useful for geometric calculations. They show invariance to transformations of position, size and viewing angle (Tanaka, et al. 1991), have access to outputs of neurons from anterior inferotemporal areas that are not transformation-invariant, appear to show rapid learning, and have been implicated in object perception (see Wallis and Bulthoff, 1999). This does not suggest that IT cells are involved in the geometric basis of object identity, but does demonstrate the existence of neurophysiological hardware with some of the appropriate characteristics.

distances or lengths ("metric information") is extracted, and has been used in hundreds of studies, such as in space perception.

Fewer studies have considered the Klein framework as a *whole*, but its importance has clearly been recognized. The philosopher Cassirer (1944, originally published 1938) was the first to explicitly apply the geometric hierarchy and groups to perception, although he credits Helmholtz and Poincare for early insights. Cassirer had the insight that the same two concepts are to be found in both modern geometry and perception: "If one surveys the facts as they have been described by psychologists, one meets again and again with two fundamental concepts that are familiar to us from another trend of thought: the concepts of 'invariance' and 'transformation'" (pp. 11). The facts he surveyed were primarily the perceptual constancies. He was concerned, however, that the systematic progression of Klein's levels towards universality would be difficult to apply to perception because geometry was open to intuition, but the perceptual constancies were not. This obstacle seems less problematic to modern thought.

In modern psychology, Mark, Todd, and Shaw (Mark and Todd, 1985; Mark, Todd, and Shaw, 1981) apply the Klein hierarchy to event perception. They suggest that the Klein classification of geometries may provide the source of information that allows observers to distinguish one type of change from another. They apply geometry to the perception of growth, and show that cardoidal strain (a topological transformation) characterizes the physical change which observers identify specifically as growth of the human head. They set out a general task: "The problem for perceptual theorists is to discover the specific variants and invariants that are perceptually salient to human observers." (pp. 856). In a different

area, Cheng (1984; Cheng and Gallistel, 1984) uses levels of the Klein hierarchy as candidates for how rats represent space in navigation. They conclude that the geometry rats use is Euclidean. Interestingly, they also conclude that only the uniquely Euclidean properties, distance and angle (along with sense), are used even though they note that logically, properties also contained in other geometries, such as cross ratio of distances, are contained in Euclidean geometry as well.   Finally, most recently, Chen, who has long argued for topology, and Todd who has used the geometry in growth, have collaborated to suggest that the levels of the hierarchy describe the salience of 3-D form properties (Todd, Chen, and Norman, 1998; see also Tittle, Todd, Perotti, and Norman, 1995). Wire forms are easiest to recognize when the distracter in a match-to-sample task is a (uniquely) topological transformation than when it is an affine transformation, which is in turn easier than when it is isometric. Chen, Todd, and colleagues, along with Bedford who uses the hierarchy for visual-motor maps (1993, 1994) and the present theory, are the among the very few researchers who have emphasized the usefulness of *multiple levels in the same observer for the same problem at the same developmental stage*.

3.2 Potential problems

The present object identity theory may appear to make predictions easily falsified by both intuition and empirical data. Consider an apparent motion experiment where there are two figures on the right hand side competing to capture a rectangle on the left hand side. It is well known that apparent motion falls off rapidly with distance. Yet, doesn't the theory predict that a very distant rectangle, otherwise identical, will capture the left hand rectangle rather than a much

nearer rectangle that has been shrunk slightly? No. Although the miniaturized rectangle is from a higher level of the hierarchy than the displaced rectangle, the amount of transformation in the two cases has been made incomparable. One prediction the theory does make is capture by the displaced rectangle if the distance of the nearest edge is no further than that of the miniature rectangle, unlike the prediction implied by other views where there should be no preference between the two rectangles. Comparisons between the levels require equating the amount of transformation from each level. Although one would hope the hierarchy were reasonably robust such that small differences are inconsequential, one would not expect a 500-fold change (for instance) in a lower level to overcome a 10% change in a higher level.

Incomparable comparisons, where a large change from a low level overcomes a small change from a higher level, are not genuine reversals of a hierarchy. An example from the traditional hierarchical interpretation of semantic memory (Collins and Quillian, 1969) may illustrate. It is faster to verify that a "canary is a bird" than a "canary is an animal", which can suggest hierarchical encoding of the concepts. Yet it may well be slower to verify that "an ostrich is a bird" than "a *canary* is an animal". This could be viewed as a reversal - a lower level category takes longer to access than a higher level. Yet here it is clearer than such a comparison would not be used as counterevidence for hierarchical structure: it is not "fair" to compare the examples because they have different typicalities. One is easy; the other hard. The requirement of equating difficulty when comparing different levels of a hierarchy is the same for the geometric hierarchy. Difficulty is determined by amount of transformation, rather than typicality of the exemplar.

Equating amount of transformation across different levels in the geometric hierarchy is challenging, but not intractable[8].

A second type of reversal that the hierarchy appears subject to involves situations where low levels of the hierarchy lead to a "different object" conclusion, but higher levels lead to a "same object" conclusion. Consider two pennies lying on a table in different places but otherwise identical. You will almost certainly judge them to be distinct objects even though they differ only in location, and thus are samples from the lowest level of the hierarchy. But now consider a bird flying past, behind a house, then emerging from the other side. In this situation, you will almost certainly identify the two temporal segments to refer to the same bird even though, due to flapping of the wings, the shapes of the two samples are the same only within the highest level of geometry. Don't these everyday situations provide counterexamples to the hierarchy?

No. What the theory does predict is that two identical pennies lying on a table in different locations will be *more* likely to refer to the same object than two pennies, one circular and one square, even if the probability of a same-object conclusion in both cases is low. Likewise, the theory predicts that two bird samples in two locations will be relatively more likely to refer to the same bird if there was no wing

---

[8]One approach is to use psychophysics of the sort used in cross-modal matching, such as when subjects are asked to adjust the brightness of light to "match" the fixed intensity of a sound. Having subjects adjust the size of a square such that it matches the fixed displacement of a target square may identify equal intensity transformations of geometric apples and oranges. Another approach when the first three levels, all linear, are to be compared is to match parameters from a general linear equation. Development of procedures to equate transformations in unclear cases will enable additional predictions of the hierarchy.

flapping, just rigid motion, then if the contours of the samples were only related topologically, even if the probability of a same-object conclusion in both cases is absolutely high. Rather than demonstrating a breakdown of the hierarchy, the examples point to the very need for the whole hierarchy. It is not the case that isometric transformations will always lead to a same object conclusion, nor that topological transformations will always lead to a different object conclusion. The outcome depends on the choices in the situation: if there is a lower level and a higher level, the lower level will prevail.

Finally, consider situations where there aren't any choices between samples from lower and higher levels. The theory has been applied to situations where a choice has to be made and the only relevant choice is between two geometries. If there is only one pair of samples - such as from the bird flying - the theory cannot make a prediction about what identity decision will be reached.. However, most everyday situations involve too many samples, not too few.

3.3. Color, pattern , and knowledge.

Everyday situations also involve choices that differ by more than just geometry. Will a red triangle capture a red circle or a green triangle? Pattern, color, texture, knowledge of the behavior of objects, and instructions are all capable of influencing the outcome of object identity. Knowledge that birds move their wings in flight may lead to the high absolute probability noted above that even shapes differing only topologically nonetheless refer to the same object in the case of birds. If a green ball disappears behind a screen and a red ball emerges from the other side, observers have the impression that there are two balls involved- more so than if the two samples were the same color. Both situations are identically geometrically, such that the

samples differ only by an isometric transformation, yet color leads to a different object judgment in one case, and a same object judgment in another. If instructions are added to inform observers that the ball changes color when heat is applied and that the screen is an electric warmer (Wilcox, 1999), it is likely that the conclusion now reverts back to "same object" despite the color changes, as the new information is incorporated in the decision.

I argue that while other factors, such as color, knowledge, and instructions can affect object identity, geometry is the more primitive, basic, and core factor in object identity. One argument comes from the developmental sequence of abilities. Wilcox (1999) has found that geometric properties can be used by infants to determine if one or two objects are present at 4.5 months of age, but pattern cannot be used for object identity until 7.5 months, and color not until infants are 11.5 months. She used an occlusion paradigm where one object goes behind one edge of a screen, and another exits the other side. (See also Baillargeon, 1994; Meltzoff and Moore, 1998; Narter, 1997; Spelke et al., 1995; Xu and Carey, 1996 for related research.) The youngest infants were surprised (longer looking times) to see the second object emerge from behind a narrow screen which was too narrow to hold two objects, but only when they were two green Styrofoam balls of different sizes, and not when they were two balls of the same size, one with dots, and the other striped. In the latter case, the infants appeared to conclude that they were one and the same ball, and could therefore "both" fit behind the narrow screen.

A second argument is that geometry can be applied generally, whereas other properties are limited in scope. For instance, color is limited to vision, whereas geometry is relevant for all modalities. Knowledge

about the properties of objects is often specific to the class of objects ("sortals" in philosophy- see Xu and Carey, 1996 for discussion). Knowledge of how birds move can be generalized to other birds, and to some extent to other animate beings, but not to rocks and tables. Short-term specialized knowledge or instructions are even more restricted; knowledge of an electric warmer that turns red objects into green ones is not very useful elsewhere. We expect geometry to be at the core of each identity problem. Special situations may be supplemented by additional relevant information. Specialized bird knowledge makes the absolute probability of a same object decision high even for topological transformations, and indispensable attributes for vision (Kubovy, see section 1.7) makes the probability of a same object decision low even for isometric transformations when two samples are presented simultaneously in vision.

Properties which constitute core knowledge may be expected in adults to be more consistent from individual to individual, to be less prone to errors, harder to modify, and less likely to require conscious access than information acquired later. Different domains may have rules relevant to identity that are unique to that domain - the theory does not preclude this - but what all the domains have in common is their ability to access a hierarchy of geometries.

3.4 Shape, "what" vs. "where", and spatial-temporal vs property/kind dichotomies

The theory leads to a surprising outcome: The properties that comprise "shape" are not fixed. Size will sometimes be a property of shape, but sometimes it will not be. Angle will sometimes characterize shape, and sometimes not. And so on. The properties that are considered shape depends

on the geometry, which changes. For instance, in Affine geometry, parallelism, straightness, and order are properties of "shape", but angle, size, and location are not. In this geometry, angle and size stand outside with location; they are not defined, and cannot distinguish between forms. That is, shape is constituted by those properties which remain unchanged by a group of transformations and are in the geometry. Properties which are altered are not in the geometry and are not shape properties- they cannot distinguish between different forms. Whether "size" or any property is a property of shape or not will actually change with geometry.

This may explain why it has been difficult to place properties in one category or the other. In a set of studies on infants and object identity, Narter (1997) puzzles over whether size should be considered a spatio-temporal characteristic ("where") or a property/kind ("what") characteristic. She notes that other researchers have viewed size as characterizing what an object is (Wilcox, 1997), but that it can also be viewed as a spatial property: "Where does size fit into this spatiotemporal versus property/kind dichotomy? ...It is possible that changes in an objects size might be more appropriately categorized as a spatial change rather than a feature change... (because) size is the amount of space an object occupies..." (pps.86-87). Similarly, if one tries to classify "orientation", there is uncertainty as to where it belongs. If the orientation of your pen changes, it is still the same pen. Like location, orientation seems not to matter to what the object is. On the other hand, if you tilt a square, it becomes something different - a diamond (Rock, 1974; see next section). Here, it looks like orientation is less like location and more like other properties that determine "shape", and some researchers seem to regard it that way (e.g. Goodale and

Milner, 1992). The present theory suggests that fixed dichotomies between spatial-temporal properties and property/kind information or between separate "what" and "where" systems (Mishkin, Ungerleider, and Mack, 1983) are problematic[9].

What it typically means to be a property of shape is whether or not that property can be used to distinguish between shapes. Whether two shapes are considered the same needs to be reconsidered. A typical description: "Two objects are the same if one can be brought into point-for-point correspondence with the other by rigid rotation and by translation. That is, if they can be superimposed, they are structurally the same; if not, they are different." (Ittelson, 1991; pp. 573). This holds true only for Euclidean geometry. Whether two shapes are the same or not depends on the geometry one is in, which changes.   A more general description: Two shapes are the same in a particular geometry if one can be brought into point-for-point correspondence with the other by any transformation in the group of transformations that defines that geometry. Determining whether or not two shapes are the same requires explicit specification of the geometry.

3.5 Extension of the theory: Orientation, Irvin Rock, and the endpoints of the hierarchy.

A classic demonstration by Irvin Rock suggests that a geometry even smaller than Euclidean may be part of the family of geometries accessed for object identity. A square rotated 90 degrees looks different; in fact, it gets a different name, a "diamond". A rotated map of Africa appears to have a different shape, and is not spontaneously

---

[9]See Goodale and Milner (1992), and Bedford (1997) for other problems with What/Where division. Also, Held's work (1970) points to what is essentially a dual sense of "where".

recognized. In Euclidean geometry, a form and a rotated form are identical, just like a form and its shifted cousin are identical. Orientation and position stand outside Euclidean geometry, where they do not contribute to shape, yet perception under object rotation suggests otherwise. Consequently, the nested hierarchy may actually be bounded by a super-small super-stringent geometry that is a subset of Euclidean geometry, within which position is still outside, but orientation no longer is. Fewer forms are equivalent, (e.g. a square and a shifted square are identical, but a square and a diamond no longer are) and the geometry is a subset of Euclidean, analogous to the way Euclidean geometry is a subset of Similarity geometry.

Rock had a quite different interpretation. He argued for the role of frames of reference. Figures get assigned "top" and "bottom", often assessed with respect to gravity. A square differs from a diamond because the top of a square is flat whereas the top of a diamond is angular. This position appears to be even more exuberantly asserted recently. According to Pinker (1998): "But as far as a geometer is concerned, they are one and the same shape. They are pegs that fit the same holes; every angle and line is the same. The only difference is how they are aligned with respect to the viewer's up-and-down reference frame, and that difference is enough to earn them different words in the English language. A square is flat on top, a diamond is pointy on top; there's no avoiding the "on top"" (pp. 266). But there is avoiding the "on top", if one dispenses with the bias that the only relevant geometry is Euclidean. The current family of geometries may make obsolete such appeals to "top" and "bottom" as causal. More elegant geometric principles, on par with understanding why squares and circles are usually perceived differently, explain why squares and diamonds can be perceived differently as well. There is a single explanatory framework of multiple nested geometries- not Euclidean geometry plus a gravitational frame of reference.

The other endpoint of shape is also noteworthy. To geometers, the properties of topology reflect the broadest and most "lenient" possible descriptions of shape. However, object identity may be able to exploit even the most radical non-topological changes when other levels of the hierarchy aren't possible, rather than stopping abruptly at the official end of geometry. In stereopsis, one point on the left retina can only match one point on the right retina (Marr, 1982), a reflection of topology as discussed earlier. Yet if there are an unequal number of samples on each eye, many-to-one matches which violate topology do occur. Panums limiting case demonstrates fusion where a line on the left can be matched to two on the right. Observers see two lines at two different distances. An expanded version where there are multiple points on one retina and twice as many on the other also leads to stereoscopic vision in which each point on the one retina gets matched to two on the other. Observers see two planes separated in depth. Real world displays which would produce such odd distributions of retinal points occur if two samples in the world just happen to line up in the line of sight for one eye, but not the other (Marr, 1982). This is clearly an improbable, though not impossible (but not possible for the expanded version). The non-topological solution as a last resort match, less preferred to topology, but not impossible, may be a sensible general reflection of reality for the most unusual, low frequency situations.

In apparent motion, topology is violated when the sample at time 1 contains

one point, but the sample at time 2 contains two. In these situations, subjects in fact report an experience of "splitting" where the first sample appears to split and go in two directions to match both time 2 samples, a known practical pitfall of the competing motion paradigm. Here too, options broader than topology to preserve object identity seem possible.

In cartoons, preservation of object identity under non-topological transformations is a source of entertainment. For instance, in the opening sequence of Warner Brothers cartoon "Animaniacs", one of the characters transforms into dozens of same-shaped miniscule pieces all of which scurry off into separate places. Geometry humor. With a nested set of geometries containing properties that range from at least orientation on the one end to even connectivity on the other, the hierarchy can be useful for almost every conceivable identity situation.

Conclusion
Object identity is pervasive. The question as to whether or not two samples refer to the exact same object must be asked regardless of whether those samples come from different times, different modalities, different eyes, or different regions of space, and regardless of whether the samples are point sources or have extended contours. Laboratory phenomena, such as apparent motion and prism adaptation, and the accomplishments they reflect, such as the detection of change, and the maintenance of accurate perceptual systems, have at their core the identical decision. How the emergence of such a general regularity can be accommodated in the current climate of ultra-modularity remains to be seen.

The resurfacing of the same problem repeatedly invites the possibility of a common solution. Geometry is abstract

enough to apply to modules of varying content. The specific geometries mirror the different transformations that occur for a single object, due both to 1) changes to the observer when perceiving the object (see constancies in section 2.3), and 2) changes to the objects of our observations themselves (e.g. thrown rocks, isometric; balloons leaking air, similarity; shadows over time, projective; biological motion, topological).

Because actual transformations can be captured by geometric transformations, it is sensible for a "mental geometry" of those transformations to be used on discrete samples to recover object identity. It works. Yet one broad geometry will not suffice: The same two samples related by the same transformation will reflect one object in one situation but two objects in a different situation. A square and a small square can be judged to refer to the exact same object, or not. The dynamic nature of the decision is reflected in a hierarchical structure of five or more geometries where a given pair of samples are considered identical in some geometries, but not in others. That the least radical transformation (and smallest geometry) in the situation be preferred for resolving object identity seems intuitively right, yet precisely why the different levels of transformations and geometries, generated by systematic removal of nested geometric properties, correspond to ordered psychological levels remains a deep mystery. How the geometry at the core of each problem interacts with domain-specific knowledge is also currently unknown.

The current work is a start towards recognition of a general problem and a possible general solution.

References

Asimov, D. (1995). There's no space like home. *The Sciences*, 35, 20-25.

Baillargeon, R., (1994). Physical reasoning in young infants: Seeking explanations for impossible events. *British Journal of Developmental Psychology*, 12, 9-33.

Bedford, F. (1993a). Perceptual learning. In D. Medin (Ed.) *The Psychology of Learning and Motivation* (Vol. 30, pp. 1-60). New York: Academic Press.

Bedford, F.L. (1993b). Perceptual and cognitive spatial learning. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 517-530.

Bedford, F.L. (1994). Of computer mice and men. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*,

Bedford, F. L. (1995) Constraints on perceptual learning: Objects and dimensions. *Cognition,* 54, 253-297.

Bedford, F. L. (1997). False categories in cognition: the Not-the-Liver Fallacy. *Cognition,* 64, 231-248.

Bedford, F. L. (1999). Keeping perception accurate. *Trends in Cognitive Sciences*, 3, 4-11.

Bedford, F. L. & Mansson, B. (Under review). Object identity, apparent motion, transformation geometry.

Bedford, F. L. & Reinke, K. S. (1993). The McCollough Effect: Dissociating retinal from spatial coordinates. *Perception and Psychophysics,* 54, 515-526.

Bergstroem, S. S. (1977). Common and relative components of reflected light as information about the illumination, colour, and three-dimensional form of objects. *Scandinavian Journal of Psychology,* 18, 180-186.

Bertelson, P. (1996) Starting from the ventriloquist: the perception of multimodal events. A*dvances in Psychological Science* (Vol 2) (Sabourin, M.,Craik, F. and Robert, M., eds), pp 419-439, Tayor & Francis.

Burt, P. & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, 88, 171-195.

Cassirer, E. (1944). The concept of group and the theory of perception. *Philosophy and Phenomenological Research*, 5, 1-36.

Chen, L. (1980). "Tolerance space and Gestalt" *Acta Psychologica* 1984, 3, 259-266.

Chen, L. (1982a). "What are the units of figure perceptual representation?" *Studies in Cognitive Science* (22) Bulletin issued by the School of Social Sciences, University of California at Irvine. Irvine, CA.

Chen, L. (1982b). "Topological structures in visual perception." *Science*. 218, 699-700.

Chen, L. (1982c). "Connectedness and the object superiority effect." *Investigative Ophthalmology and Visual Science, Supplement*. 22, 124.

Chen, L. (1985). Topological structure in apparent motion. *Perception*, 14, 197-208.

Cheng, K. (1984). The primacy of metric properties in the rat's sense of place. Dissertation, University of Pennsylvania, *Dissertation Abstracts International*, 45, 07B, 2338.

Cheng, K. & Gallistel, C.R. (1984). Testing the geometric power of an animal's spatial representation. In H. Roitblat, T.G. Bever, & H. Terrace (Eds.), *Animal Cognition* (pp. 409-423). Hillsdale, NJ: Lawrence Erlbaum.

Choe, C.H., Welch, R.B., Gilford, R.M., & Joula J.F.(1975). The "ventriloquist effect": visual dominance or response bias?. *Perception and Psychophysics*, 18, 55-60.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior,* 8, 240-247.

Cunningham, H.A. (1984). An Apple microcomputer-based laboratory for the study of visual-motor behavior. *Behavior Research Methods, Instrument*s, and Computers, 17, 484-488.

Cutting, J (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 323-334.

Cutting, J.E. (1988). Affine distortions of pictorial space: some predictions for Goldstien (1987) that La Gournerie (1859) might have made. *Journal of Experimental Psychology: Human Perception and Performance,* 14 (2), 305-311.

Dawson, M. R. (1989). Apparent motion and element connectedness. Spatial Vision, 7, 241-251.

Dawkins, R. (1996) *Climbing Mount Improbable*. London: Viking.

Dennett, D.C. (1996). *Darwin's Dangerous Idea : Evolution and the Meanings of Life.* Touchstone Books.

Epstein, W. (1975). Recalibration by pairing: A process of perceptual learning. *Perception*, 4, 59-72.

Farrell, J.E. & Shepard, R.N. (1981). Shape orientation, and apparent rotation motion. *Journal of Experimental Psychology:*
Human Perception and Performance, 7, 477-486.

Forster, K.I. & Taft, M. (1994) Bodies, antibodies, and neighborhood-density effects in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 844-863.

Frisby, J.P. & Clatworthy, J.L. (1975). Learning to see complex random dot stereograms. *Perception*, 4, 173-178.

Goodale, M. A & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in neuroscience*, *15*, 20-25.

Green, M. & Odom, J. V (1986). Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation. *Science*, 233, 1427-1429.

Harris, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychological Review*, 72, 419-444.

Held, R. (1970). Two modes of processing spatially distributed information. In F.O. Schmitt (Ed.) *The neurosciences second study program.* New York: Rockefeller University Press.

Helmholtz, H.E.F. von (1962). *Treatise on Physiological Optics.* Southall, J.P.C., ed./trans. Dover (Originally published 1909).

Hering, E. (1905) Grundzuge der Lehre vom Litchtsinn. In Handbuch der gesamter Augenheilkunde, vol 3, chap. 13. Berlin. see Woodworth (1938.

Ittelson, W.H., Mowafy, L. & Magid, D. (1991). The perception of mirror-reflected objects, *Perception*, 20, 567-584.

Jack, C.E. & Thurlow, W.R. (1973). Effects of degree of visual association and angle of displacement on the ventriloquism effect. *Perceptual and Motor Skills,* 38, 967-979.

Jackson, C.V. (1953). Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, 5, 52-65.

Johansson, G. (1974) Visual perception of rotary motion as transformations of conic sections: A contribution to the theory of visual space perception. *Psychologia: An International Journal of Psychology in the Orient,* 17, 226-237.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago: University of Chicago Press.

Kahneman, D. & Treisman, A. & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. Cognitive Psychology, 24, 175-219.

Klein, F. (1957). Vorlesungen uber hohere geometrie (Lectures on higher geometry) (3rd ed.) New York: Chelsea. (Original work published 1893)

Kolers, P. (1972). *Aspects of motion perception*. Oxford, England: Pergamon Press

Kolers, P.A. & Pomerantz, J.R. (1971). Figural change in apparent motion. *Journal of Experimental Psychology*, 87, 99-108.

Kubovy, M. (1981). Concurrent-Pitch segregation and the theory of indispensable attributes., In M. Kubovy & J.R. Pomerantz (eds.), *Percpetual Organization*, Erlbaum: Hillsdale, NJ., 55-96.

Kubovy, M. (1988). *Journal of Experimental Psychology: Human Perception & Performance*, 14, 318-320.

Leslie, A.M., Xu, F., Tremoulet, P.D. & Scholl, B.J. (1998). Indexing and the object concept: developing 'what' and 'where' systems. *Trends in Cognitive Sciences*, 2, 10-18.

Leyton, M. (1992). *Symmetry, Causality, Mind. Cambridge, Mass: MIT Press.*

Mack, A., Klein, L., Hill, J. & Palumbo, D. (1989). Apparent motion: Evidence of the influence of shape, slant, and size on the correspondence process. *Perception and Psychophysics*, 46, 201-206.

Mark, L.S., Todd, J.T., & Shaw, R.E. (1981). Perception of growth: A geometric analysis of how different styles of change are distinguished. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 855-868.

Mark, L.S. & Todd, J.T. (1985). Describing perceptual information about human growth in terms of geometric invariants. *Perception and Psychophysics*, 37 (3), 249-256.

Marr, D. (1982). *Vision*. New York, NY: W.H. Freeman and Co.

Medin, D.L., Goldstone, R.L. & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100, 254-278.

Meltzoff, A.N. & Moore M.K. (1998). Object representation, identity, and the paradox of early permanence: Steps toward a new framework. *Infant Behavior & Development*, 21 (2), 201-235.

Mishkin, M., Ungerleider, L. G. & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, *6*, 414-417.

Narter, D.B. (1997). Infants' expectations about the spatial and physical properties of a hidden object. Dissertation. University of Arizona.

Navon, D. (1976). Irrelevance of figural identity for resolving ambiguities in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 130-138.

Niall, K. K. & MacNamara, J. (1989). Projective invariance and visual shape constancy. *Acta Psychologica*, 72, 65-79.

Niall, K. K. & MacNamara, J. (1990). Projective invariance and picture perception. *Perception*, 19, 637-660.

Olson, R.K. (1974). Slant judgments from static and rotating trapezoids correspond to rules of perspective geometry. Perception & Psychophysics, 15, 509-516.

Orlansky, J. (1940). The effect of similarity and difference in form on apparent visual movement. *Archives of Psychology*, 246, 85.

Perkins, D.N. (1976). How good a bet is good form? *Perception*, 5, 393-406.

Pinker, S. (1997). *How the mind works*. W.W. Norton & Co.

Radeau, M (1994). Auditory-visual spatial interaction and modularity. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 13, 3-51.

Radeau, M. & Bertelson, P. (1974). The aftereffects of ventriloquism. *Quarterly Journal of Experimental Psychology,* 26, 63-71.

Radeau, M. & Bertelson, P. (1976). The effect of a textured visual field on modality dominance in a ventriloquism situation. *Perception and Psychophysics,* 20, 227-235.

Radeau, M. & Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations, *Perception and Psychophysics*, 22, 137-146.

Radeau, M. & Bertelson, P. (1987). Auditory-visual interaction and the timing of inputs, Thomas (1941) revisited. *Psychological Research*. 49, 17-22.

Redding, G.M. & Wallace, B. (1976). Components of displacement adaptation in acquisition and decay as a function of hand and hall exposure. *Perception and Psychophysics*, 20, 453-459.

Rescorla, R.A. (1985). Pavlovian conditioning analogues to Gestalt perceptual principles. In F.R. Brush & J.B. Overmier (Eds.), *Affect, conditioning and cognition: Essays on the determinants of behavior*. Hillsdale NJ: Erlbaum.

Rock, I. (1983). The logic of perception. Cambridge, Mass: Bradford Books and MIT Press.

Rock, I. (1974). The perception of disoriented figures. *Scientific American*, 230, 78-85.

Rosser, R.A., Narter, D.B., & Paullette, K.M. (1995). *Ontological kind, object identity, and infants' sensitivity to violations of the continuity constraint*. Poster presented at the meeting of the Society for Research in Child Development, Indianapolis, IN>

Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J.A. Sprague & A.N. Epstein (Eds.), *Progress in psychobiology and physiological psychology* (Vol. 6) New York: Academic Press.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, *13*, 501-518.

Schacter, D. L. (1992). Understanding implicit memory: A cognitive neuroscience approach. *American Psychologist*, *47*, 559-569.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge [England]; New York : Cambridge University Press.

Shepard, R.N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, 91, 417-447.

Shepard, R.N. (2001). Perceptual-cognitive universals as reflections of the world. *Brain and Behavioral Sciences, 24.*

Shepard, R.N. & Judd, S.A. (1976). Perceptual illusion of rotation of three-dimensional objects. *Science*, 191, 952-954.

Sigman, E. & Rock, I (1974) Stroboscopic movement based on perceptual intelligence. *Perception, 3*, 9-28.

Spelke, E.S. et al. (1995) Spatiotemporal continuity, smoothness of motion and

object identity in infancy *British Journal of Developmental. Psychology*. 13, 113-142

Spelke, E.S., Kestenbaum, R. (1986). Les origines du concept d'objet. *Psychologie Francaise*, 31, 67-72.

Stratton, (1897). Perception without inversion of the retinal image. *Psychological Review*

Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. London: Methuen.

Tanaka, K., Siato, H.A., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in inferotemporal cortex of the Macaque monkey. *Journal of Neurophysiology*, 66, 170-189.

Thomas, G.J. (1941). Experimental study of the influence of vision on sound location. *Journal of Experimental Psychology*, 28, 167-177.

Todd, J.T., Chen, L., & Norman J.F. (1998). On the relative salience of Euclidean, affine, and topological structure for 3-D form discrimination. *Perception*, 27, 273-282.

Tittle, J.S., Todd, J.T., Perotti, V.J. & Norman, J.F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 663-678.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Wallis, G. & Bulthoff, H. (1999). Learning to recognize objects. *Trends in Cognitive Sciences*, *3*, 22-30.

Warren, W.H. (1977). Visual information for object identity in apparent movement. *Perception and Psychophysics,* 25, 205-208.

Warren, D.H., Welch, R.B. & McCarthy, T. J. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception and Psychophysics,* 30, 557-564.

Weiss (1941). Self-differentiation of the basic patterns of coordination. *Comparitive Psychology Monograph*, 17, 1-96.

Welch, R.B. (1986). Adaptation of space perception, in *The Handbook of Perception and Performance* (Vol. 1)(Boff, K. Kaufman, L. and J.P. Thomas, J.P., eds), pp. 24.1-24.45, Wiley

Welch, R.B. (1978). *Perceptual Modification*. New York: Academic Press.

Welch, R.B. (1972). The effect of experienced limb identity upon adaptation to stimulated displacement of the visual field., 12, 453-456.

Welch, R.B. & Warren D.H. (1980) Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638-667.

Wertheimer, M. (1923). Untersuchgen zur lehre von der Gestalt, II. *Psychological Forshung*, 4, 301-350.

Wilcox, T. (1997). *4.5 and 7.5-month-old infants' use of shape, color, and size when reasoning about object identity*. Poster presented at the biennial meeting of the Society for Research in child Development, Washington, DC.

Wilcox, T. (1999) Object individuation: Infants' use of shape, size, pattern, and color. *Cognition*, 72, 125-166.

Wilcox, T., Nadel, L., & Rosser, R. (1996). Location memory in healthy preterm and full-term infants. *Infant Behavior and development*,19, 309-323.

Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.

Figure Captions

Figure 1. Five geometries. Schematic of nesting of the geometries. Euclidean geometry is a special case or subset of the more general Similarity geometry. Similarity geometry is a subset of Affine geometry etc. Callouts show some of the properties contained within each geometry.

Figure 2. Geometric transformations - 2 dimensions.. Sample forms before and after sample transformations within each group of transformations.

Figure 3. Geometric transformations - 1 dimension. A line before and after sample transformations within each group of transformations. Similarity, Affine, and Projective collapse to the same level for 1 dimensional "forms" (i.e. lines).

Figure 4.   Spatial-temporal regularities. Schematic of simple apparent motion situations and Korte's third law. Space (S) is shown on the X axis and time (T) on the Y axis. Each quadrilateral shows 1 pair of dots in 1 apparent motion situation: A point on the graph shows the location in time and space of a punctate visual stimulus and dotted lines show the changes in space and time. Square 1a-2a shows an original apparent motion situation. Squares 1b-2b (3 of them) show isometric transformations of that situations. Square 1c-2c shows a similarity transformation. Square 1d-2d shows an affine transformation.

Appendix 1

A. The groups of transformations in equation form ordered from most general (biggest) to most specific (smallest). For requirements to be a group and application of groups to perception see, e.g. Cassirer, 1944 and Mark et al., 1981)

Topological
$X' = f(X,Y)$
$Y' = g(X,Y)$

where f and g are continuous functions
(if f and g are the specific functions below, then the subset is projective)

Projective
$X' = (a1X + a2Y + a3)/(c1X + c2Y + c3)$
$Y' = (b1X + b2Y + b3)/(c1X + c2Y + c3)$
where determinate $\neq 0$
(if c1=0 and c2=d, then this reduces to affine)

Affine
$X' = aX + bY + m$
$Y' = cX + dY + n$
where $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$
Note these are the most general linear equations.
(if c= - b and d=a then this reduces to similarity)

Similarity
$X' = aX + bY + m$
$Y' = -bX + aY + n$ or
$X' = aX + bY + m$, $Y' = bX - aY + n$
(if $a^2 + b^2 = 1$ then this reduces to isometric)

Isometric
$X' = aX + bY + m$
$Y' = -bX + aY + n$
where $a^2 + b^2 = 1$

B. Properties altered and preserved by the 5 different groups of transformations in the order from the least to most radical.

Isometric.
Properties altered: position
Properties preserved: lots - e.g. distance between any 2 points (length), collinear points ("in a line") remain collinear and non-collinear points remain non-collinear, parallel lines remain parallel, angle, order of points

Similarity.
altered: distances between any 2 points, position
preserved: ratio of distances between any 2 pairs of points (A'B')/(C'D') = AB/CD, collinearity, parallelism, angle, order

Affine
altered: angle, area, distances between any 2 points, position
preserved: ratio of distances between points along the same line (or parallel lines), collinearity, parallelism, order.
 Other useful properties preserved are being a conic, and the degree of equation. That is one type of conic section will stay the same type of conic after an affine transformation (parabola stays a parabola, ellipse stays an ellipse, hyperbola stays a hyperbola). An example of degree of equation is that a cubic stays a cubic.

Projective
altered: parallelism, angle, area, distances between any 2 points, position, being a type of conic.
preserved: collinearity (provided the mapping is onto), order. Also *cross* ratio, degree of curve.

Topological
altered: collinearity, parallelism, angle, area, distances between any 2 points, position, being a type of conic.
preserved: order. Also connectivity, being a closed curve (closed curves get mapped to closed curves), endlessness of a curve, incidence relations, cyclic order.

Appendix 2 Spatial-temporal properties and Korte's third law.

The geometries have been applied *separately* to space and time; an extension of the object identity theory is to combine space and time into the same multidimensional "space". When the samples are point sources rather than extended contours, complex geometric transformations are eliminated. Changes of idealized point sources in location and time become more salient and what are known as "spatial-temporal properties" emerge. Classic spatial-temporal laws can be considered a subset of the more general geometric laws. For instance, Korte's third law is a spatial-temporal regularity that occurs in apparent motion; if the spatial separation between two stimuli increases, then the time difference between the two stimuli must also increase in order to maintain optimal apparent motion. The law can be rederived from geometry.

When this regularity was observed, apparent motion experiments were conducted using small and unchanging stimuli that had little spatial extent/form and instead could be characterized as point sources. When the two samples are individual points rather than extended contours, the number of geometric transformations which are possible become greatly reduced. A point in space can only be moved to a new location (isometric) or it can be split into two points (non-topological). Similarly, a point in time can be moved to a new time, or split into two events. Stimuli which have limited and unchanging duration from one sample to the next can be viewed as a point sources for time, also like many apparent motion situations. The simple two-dimensional apparent motion case where time is one dimension, space is also one dimension (i.e. the stimulus at time 1 is constrained to change position only along the horizontal for time 2), and the stimuli are

points sources in space and time, is shown in Figure 4. The samples at time 1 and time 2 are localized on the graph as points and the change in space and time are shown with dotted lines. Thus, each apparent motion situation is represented as a square. Due to the restrictions on keeping the extent of the stimulus the same in space and in time, what remains are the following geometric transformations.

If the square [shown as 1a-2a] is picked up and moved elsewhere, that is an isometric transformation which should be largely innocuous. That is, if the apparent motion displays are moved to a different part of the room, this change should have minimal effect on the quality of apparent motion (translations along x); if the experiment is begun early or later, that too should not affect the strength of the phenomenon (translations along y), nor should any combination of the two translations [the 3 isometric transformations shown as the 3 squares 1b-2b]. These transformations are typically not conducted as experimental manipulations because they are expected to be largely uninfluential, as they are in the geometric analysis as well. If the square is stretched, then the different conditions that constitute Korte's third law can be seen. If the spatial separation is increased, but not the temporal separation [shown as the rectangle 1d-2d], this should disrupt optimal apparent motion more than if both separations are increased (equally) [shown as square 1c-2c]. This is what is expected; the former is an affine transformation (the square became a rectangle) which is a more radical transformation and less like the original than the latter transformation, which is a similarity transformation (the square became a different size square). Korte's law doesn't specify that space and time be altered *equally* but this presumably reflects the difficulty of independently

equating units of space and time. The closer one can get to equal shifts, the better optimal motion will be maintained under spatial and temporal separation shifts.

This example demonstrates that a classic law can be considered a subset of this more general framework. In addition, it shows the geometric hierarchy can encompass the limiting cases of points rather than forms and provides a framework for understanding "spatial-temporal" properties as well as the more "form" based ones. Finally, it also points to extensions of the hierarchy where time and space are analyzed in the same hierarchy.