Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension ☆

Lara L. Jones [a,*], Zachary Estes [b]

[a] Department of Psychology, University of Georgia, Athens, GA 30602, USA
[b] Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom

Abstract

Bowdle and Gentner (2005) proposed a reconciliation of the comparison and categorization models of metaphor comprehension. Their career of metaphor model posits that, as a metaphorical term becomes more conventional, its mode of processing shifts from comparison to categorization. However, other recent studies (Chiappe, Kennedy, & Chiappe, 2003; Jones & Estes, 2005) suggest instead that aptness may mediate metaphorical processing. We empirically contrasted conventionality and aptness to examine their roles in metaphor comprehension. Aptness predicted the preference for metaphors over similes (Experiment 1), the speed and ease of metaphor comprehension (Experiment 2), and the category membership of metaphorical terms (Experiment 3). Conventionality did not reliably predict any of these aspects of metaphorical processing. Thus, results supported the categorization model, and failed to support the career of metaphor.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Metaphor comprehension; Career of metaphor; Conventionality; Aptness

The comprehension of nominal metaphors (e.g., *Discipline is fertilizer*) has been explained by two contrasting modes of processing. Comparison models (e.g., Gentner & Wolff, 1997; Miller, 1993) posit that metaphors are understood as analogies: the properties of the metaphor vehicle (e.g., FERTILIZER promotes growth) are aligned with and then projected onto the metaphor topic (e.g., DISCIPLINE). Alternatively, categorization models (Glucksberg, 2003; Glucksberg & Keysar, 1990; Glucksberg, McGlone, & Manfredi, 1997) claim that metaphors are understood as categorical assertions: the topic (e.g., DISCIPLINE) is included in the category of the vehicle (e.g., FERTILIZER), and therefore inherits relevant properties of that category (e.g., promotes growth). Bowdle and Gentner (2005) recently proposed a hybrid of these contrasting models. According to their *career of metaphor* model, the processing of any given metaphor is determined by its conventionality—that is, the strength of association between the vehicle (e.g., FERTILIZER) and its figurative property (e.g., promotes growth). Specifically, as a metaphor vehicle becomes conventionalized across its "career" of usage, it switches from com-

* Corresponding author.
  E-mail address: laraj@uga.edu (L.L. Jones).

parative to categorical processing. So metaphors with a novel vehicle are understood by comparison, whereas metaphors with a conventional vehicle may be understood by categorization.

In this article we re-examine whether conventionality predicts metaphor comprehension, as claimed by the career model of metaphor. We argue instead that metaphor processing is mediated by aptness—that is, the degree to which the figurative meaning of the vehicle describes a relevant feature of the topic. Although conventionality and aptness are theoretically independent factors (Chiappe et al., 2003), a brief review of the literature indicates that the two factors exhibit strikingly similar influences on metaphorical processing. We suggest that this is no coincidence. Rather, studies of conventionality and studies of aptness have produced similar results because researchers have routinely confounded the two factors. This empirical shortcoming is theoretically critical: whether conventionality or aptness truly mediates metaphorical processing will discriminate between the categorization model and the career model of metaphor. We therefore report three experiments in which conventionality and aptness were manipulated independently. First though we define conventionality and aptness, and then we outline their hypothesized roles in the different models of metaphor comprehension.

Conventionality and aptness

Conventionality is the strength of association between a metaphor vehicle and its figurative meaning (Bowdle & Gentner, 2005; Gentner & Wolff, 1997; Wolff & Gentner, 2000; see also Giora, 1997).[1] Some terms (e.g., ROLLER COASTER) are frequently used in a metaphorical sense (e.g., *Their marriage was a roller coaster*), and hence they become associated with their figurative meaning (i.e., having highs and lows). Less conventional terms (e.g., RAIL), in contrast, have no salient figurative meaning as a consequence of having little metaphorical usage. So the more a term is used metaphorically, the more conventional it becomes.

Aptness is the extent to which the vehicle's figurative meaning expresses an important feature of the topic (Blasko & Connine, 1993; Chiappe & Kennedy, 1999; Chiappe et al., 2003; Gerrig & Healy, 1983; Glucksberg & McGlone, 1999). For a metaphor to be apt, two conditions must be met. First, the vehicle term must have a

salient property for attribution. For some vehicles, that property is strongly associated with the term (i.e., the vehicle is conventional). But novel vehicles may also produce apt metaphors, despite their relatively weak association between the term and its attributive property (Camac & Glucksberg, 1984). For example, RAIL may be a relatively novel vehicle, but the metaphor *That fashion model is a rail* is nevertheless highly apt because the property implied by the vehicle (i.e., extremely thin) is an important feature of fashion models. The second necessary condition of aptness, then, is that the salient property of the vehicle must be relevant to the topic. To borrow an example from Glucksberg and McGlone (1999, p. 1544), the vehicle BUTCHER has a salient metaphorical meaning (i.e., ''bungling''). But some topics (e.g., PIANISTS) can more easily accept that property than can others (e.g., FILE CLERKS). If the property implied by the vehicle is irrelevant to the topic, then the metaphor will be less apt (e.g., *My filing clerk is a butcher*). Thus it is the interaction between topic and vehicle that is critical for aptness.[2]

Models of metaphor comprehension

The various models of metaphor comprehension are reviewed extensively elsewhere (e.g., Bowdle & Gentner, 2005; Glucksberg, 2003). Here we provide only a brief description of each model, with emphasis on the hypothesized roles of conventionality and aptness. Comparison models posit that metaphor comprehension entails an alignment process followed by a projection process (e.g., Gentner, 1983; Gentner & Wolff, 1997; Miller, 1993). During the alignment stage, the topic and vehicle concepts perform identical roles as the subjects of a symmetric comparison process. But the second stage is directional, with the vehicle properties being projected onto the topic (Gentner & Wolff, 1997; Wolff & Gentner, 2000). Using the metaphor *Men are wolves* as an example, the alignment process first identifies the shared relation ''prey on'' between men and wolves. Next, the non-identical arguments of this relation are aligned (i.e., MEN with WOLVES, and WOMEN with ANIMALS). Finally, unique vehicle properties connected to the shared relation (e.g., the instinctive cause of predation) are projected onto the topic. Thus, the metaphor is compre-

---

[1] For brevity, we restrict our use of the term ''conventionality'' to refer only to the association between the vehicle and its metaphorical property. This usage differs from studies in which conventionality refers to the familiarity of a metaphor's topic-vehicle pairing (e.g., *Love is a journey;* Gibbs, 1992; Lakoff & Johnson, 1980).

[2] A related manner in which aptness differs from conventionality is that, unlike conventionality, aptness describes the relative position of the topic and vehicle within their respective domains (Tourangeau & Sternberg, 1981, 1982). For example, *Some robins are lions* is less apt than *Some eagles are lions*, because both EAGLES and LIONS are similarly predatory, whereas ROBINS are not. Conventionality, in contrast, refers only to the vehicle of the metaphor, and therefore is independent of its topic.

hended as an analogy: men prey on women just as wolves prey on other animals. Bowdle and Gentner (2005) explain that "distinctive elements of the base [i.e., vehicle] can be transferred to the target [i.e., topic] to the extent that they are connected to the common relational system" (p. 197; see also Gentner & Wolff, 1997, p. 335; Wolff & Gentner, 2000, p. 529). It follows that metaphor comprehension should be facilitated by any factor that emphasizes (1) distinctive elements of the vehicle, or (2) the common relational system. Gentner and Wolff (1997; Wolff and Gentner, 2000) showed that comprehension indeed was facilitated by vehicle conventionality and relational similarity which, respectively, support predictions (1) and (2). But notice that, by this model, the topic plays only a passive role in comprehension—once the relational alignment has occurred, it is the vehicle that determines the properties to be projected. Thus it also follows that aptness should not be a primary predictor of metaphor comprehension, since aptness relies on both the topic and the vehicle (see Bowdle & Gentner, 2005, p. 204–205).

According to the categorization model, metaphor comprehension entails the inclusion of the topic as a member of the figurative category named by the vehicle (Glucksberg, 2003; Glucksberg & Keysar, 1990; Glucksberg et al., 1997). Using the previous metaphor, the vehicle refers not to the literal category of WOLVES, but rather to the abstract figurative category of PREDATORY CREATURES. And because the topic (i.e., MEN) is understood as a member of that figurative category, the topic therefore inherits the relevant features of the category (e.g., instinctive predation). So the vehicle of a metaphor has dual reference: it refers to an abstract figurative category as well as its own literal category (Glucksberg et al., 1997; McGlone, 1996; McGlone & Manfredi, 2001). By this model, the topic and vehicle concepts perform different functions from the outset of the comprehension process. Specifically, the vehicle provides salient properties for attribution, whereas the topic constrains which dimensions are relevant for attribution (Glucksberg & McGlone, 1999; Glucksberg et al., 1997). It follows from this view of metaphor as the interaction of topic and vehicle (see also Black, 1962, 1993) that conventionality should not be a primary predictor of metaphor comprehension, because conventionality refers to only the vehicle concept (Bowdle & Gentner, 2005; Gentner & Wolff, 1997; Wolff & Gentner, 2000). Instead, aptness should predict comprehension, because aptness reflects both the salience of the vehicle property and its relevance to the topic: "Metaphors will be easily understood when the newly created classification is perceived as relevant and informative...Whether a [categorization] is relevant and informative will depend, of course, on...whether the metaphor vehicle has salient properties that are diagnostic and relevant to that topic" (Glucksberg & Keysar, 1990, p. 13; see also Bowdle and

Gentner, pp. 204–205; Glucksberg & McGlone, 1999, p. 1544; Glucksberg et al., 1997, pp. 57–59; McGlone & Manfredi, 2001, p. 1210). Thus, the categorization model predicts an effect of aptness, but no effect of conventionality.

Finally, the career of metaphor (Bowdle & Gentner, 1999, 2005; Gentner & Bowdle, 2001; Gentner, Bowdle, Wolff, & Boronat, 2001; Gentner & Wolff, 1997; Wolff & Gentner, 2000) maintains that vehicle conventionality produces a shift from comparative processing for novel metaphors (e.g., *Science is a glacier*) to categorical processing for conventional metaphors (e.g., *Genes are blueprints*). This shift in processing is a byproduct of the metaphor vehicle acquiring dual reference. Initially, a vehicle has only its literal reference, but as that vehicle gains in metaphorical usage across time, it may eventually acquire a figurative reference as well. Thus, different concepts may be more or less strongly associated with their figurative reference (cf. Giora, 1997), and the strength of this association determines whether comparison or categorization will occur. Like the comparison model, the career model also predicts no effect of aptness on metaphor comprehension (see Bowdle & Gentner, 2005, pp. 204–205), because the topic concept is assumed to play no active role in property projection. Instead, metaphor comprehension should be mediated by conventionality. Thus the career model and the categorization model make opposing predictions with regard to aptness and conventionality.

Confounding of conventionality and aptness

As previously described, the categorization model emphasizes the role of aptness in metaphor comprehension, whereas the career model emphasizes conventionality. In this section we consider which of these two factors is more likely to mediate metaphorical processing. Researchers have used a number of tasks to investigate the influence of conventionality and/or aptness on metaphor comprehension. The dependent variables from these tasks are listed in Table 1, along with relevant results from several studies. As summarized in the table, aptness correlates with each of these measures.

In the "form preference" task (e.g., Aisenmann, 1999; Gibb & Wales, 1990), participants indicate whether they prefer a given statement as a metaphor (e.g., *Discipline is fertilizer*) or as a simile (e.g., *Discipline is like fertilizer*). According to the career model, novel vehicles induce comparative processing. So given that similes have the form of a comparative statement (i.e., *X is like a Y*), novel statements should be preferred as similes. Conventional vehicles, in contrast, may elicit categorical processing. So because metaphors have the form of a categorical statement (i.e., *X is a Y*), conventional statements should be preferred as metaphors.

Table 1
Summary of correlations between aptness and measures of metaphor comprehension

Dependent Variable	Correlation	Source
Metaphor preference	$r = -.65$, $p < .01$[a]	Bowdle and Gentner (2005), Experiment 1
	$r = +.75$, $p < .01$	Chiappe and Kennedy (1999)
	$r = +.61$, $p < .01$	Chiappe and Kennedy (2001), Experiment 1
	$r = +.75$, $p < .01$	Chiappe and Kennedy (2001), Experiment 2
	$r = +.75$, $p < .001$	Chiappe et al. (2003)
	$r = +.63$, $p < .001$	Chiappe et al. (2003)
	$r = +.45$, $p < .05$	Utsumi and Kuwabara (2005)
Comprehension latencies	$r = -.13$, $p = .64$[a,b,c]	Bowdle and Gentner (2005), Experiment 2
	$r = -.55$, $p < .05$[b]	Chiappe et al. (2003)
	$r = -.46$, $p < .01$[b]	Gagné (2002), Experiment 1
Comprehensibility ratings	$r = +.86$, $p < .001$[b]	Chiappe et al. (2003)
	$r = +.81$, $p < .001$[b]	Gagné (2002), Experiment 1
	$r = +.81$	Katz et al. (1988), Table 3
	$r = +.84$, $p < .01$	Kusumi (1987), Table 2
	$r = +.83$	Sternberg and Nigro, Table 2, Form 4
	$r > +.90$	Tourangeau & Rips, Experiment 1
	$r = +.64$, $p < .01$	Tourangeau and Sternberg (1981), Table 6
Metaphorical categorization	$r = +.75$, $p < .001$	Jones and Estes (2005), Experiment 2

[a] Novel items only.
[b] Aptness ratings averaged across metaphors and similes.
[c] Difference in latencies between metaphor and simile forms.

Indeed, several studies have shown that the preference for the metaphorical form of a statement is greater when that statement has a conventional vehicle than when it has a novel vehicle (Bowdle & Gentner, 1999, 2005; Gentner & Bowdle, 2001; Nakamoto & Kusumi, 2004; Zharikov & Gentner, 2002).

Although this result initially appears to support the career model of metaphor, it must be noted that those studies did not control for aptness. In fact, Bowdle and Gentner (2005) discovered in a post-test that their conventional items ($M = 6.54$ out of 10, $SD = 1.27$) were more apt than their novel items ($M = 4.18$, $SD = 1.68$), thus indicating that they had inadvertently confounded conventionality and aptness. Moreover, as shown in Table 1, several other studies have found that aptness predicts metaphor preference (Chiappe & Kennedy, 1999, 2001; Chiappe et al., 2003; Chiappe, Kennedy, & Smykowski, 2003; Utsumi & Kuwabara, 2005). Thus, it remains unclear which is the true mediator of metaphor preference. In the most direct comparison of these two factors, Chiappe et al. (2003) collected form preference ratings, conventionality ratings, and aptness ratings for the same set of metaphors and similes. They obtained a significant positive correlation between aptness and metaphor preference ($r = +.63$; see Table 1), but no relation between conventionality and metaphor preference ($r = +.01$).

Another common measure of metaphor processing is ease-of-comprehension, as indicated by comprehension latencies and comprehensibility ratings. On average, con-

ventional metaphors are comprehended faster than novel metaphors (Blank, 1988; Bowdle & Gentner, 2005; Gentner & Wolff, 1997). Again, these studies *may* support the career model of metaphor, but because they confounded aptness with conventionality (see Bowdle & Gentner, 2005, pp. 204–205; Jones & Estes, 2005, p. 118), the result is equivocal. This concern is validated by the finding that highly apt statements are also comprehended faster (Chiappe et al., 2003; Gagné, 2002) and easier (Chiappe et al., 2003; Gagné, 2002; Kusumi, 1987; Sternberg & Nigro, 1983; Tourangeau & Rips, 1991; Tourangeau & Sternberg, 1981) than less apt statements (see Table 1). But unfortunately these studies suffer from the converse problem—in manipulating aptness, they failed to control for conventionality. Collectively, then, these studies do not clarify whether comprehensibility is predicted by aptness or by conventionality.

To directly test the categorization model, we (Jones & Estes, 2005) introduced a categorization task to measure metaphorical processing. Using the stimuli of Gentner and Wolff (1997), in our first two experiments we found that the topic of a metaphor was more likely to be included in the vehicle category when that metaphor was conventional than when it was novel. For example, after reading the metaphor *His desk is a junkyard*, participants were relatively likely to judge that a desk could belong in the category JUNKYARD. However, in a post-test we also discovered that conventionality and aptness were confounded in that item set. We then re-analyzed our data, but with aptness as a covariate, and the effect

of conventionality disappeared. In a third experiment we manipulated aptness instead of conventionality, and we obtained a significant influence of aptness on metaphorical categorization. Unfortunately, however, aptness and conventionality were not contrasted in that experiment. So although this study suggests that aptness rather than conventionality predicts metaphorical categorization, it is ultimately no more conclusive than the studies described above.

In sum, it remains unclear whether metaphor processing is mediated by conventionality or by aptness. The overwhelming majority of studies have investigated either factor without controlling the other. As a result, those studies may well have confounded the two factors. Of the few studies that have directly investigated both conventionality and aptness, two of those analyses (Bowdle & Gentner, 2005; Jones & Estes, 2005) were post hoc and therefore not well controlled, and the other analysis (Chiappe et al., 2003) was a priori but correlational. To reiterate, our concern is not merely methodological. Rather, determining the true mediator of metaphorical processing will critically discriminate between models of metaphor comprehension: the career model of Bowdle and Gentner (2005) assumes conventionality as the mediator, whereas the categorization model of Glucksberg and colleagues (Glucksberg, 2003; Glucksberg & McGlone, 1999; Glucksberg et al., 1997; McGlone & Manfredi, 2001) suggests aptness as the mediator. We therefore conducted a series of experiments in which these two factors were manipulated independently.

Overview of experiments

In the present experiments we manipulated conventionality by varying the vehicle concepts—some vehicles were novel and some were conventional. We manipulated aptness by varying the topic. Because aptness takes into account the topic and the vehicle, metaphors having the same vehicle can vary in aptness. For instance, the metaphor *Dancers can be butterflies* is more apt than *Soccer players can be butterflies*, because the attributed property (i.e., graceful) is more typical of DANCERS than of SOCCER PLAYERS. Thus, aptness and conventionality were independently manipulated by varying the topic and vehicle concepts, respectively.

We conducted three experiments, each using a different paradigm, so as to provide convergent evidence of the roles of conventionality and aptness in metaphor comprehension. All four of the dependent variables summarized in Table 1 were used in the present experiments. Experiment 1 used the form preference task, and was directly modeled after Bowdle and Gentner's (2005) Experiment 1. In our Experiment 2, which was modeled after Bowdle and Gentner's (2005) Experiment

2, we measured comprehension latencies and collected comprehensibility ratings. Finally, in Experiment 3, we used the metaphorical categorization task devised by Jones and Estes (2005).

Experiment 1

Participants indicated whether (and the extent to which) they preferred a given statement in metaphor form (e.g., *Education is a lantern*) or simile form (e.g., *Education is like a lantern*). According to the career model, conventional statements may be processed categorically. So because metaphors have a categorical structure (*X is a Y*), it follows that conventional statements should be preferred in metaphor form. Conversely, because novel statements are processed comparatively, and because similes have a comparative structure (*X is like a Y*), it follows that novel statements should be preferred in simile form. Thus, the career model predicts a main effect of conventionality, such that preference for the metaphor form is greater for conventional than for novel statements (Bowdle & Gentner, 2005). Alternatively, the categorization model predicts a main effect of aptness (Glucksberg & McGlone, 1999). As aptness decreases, the likelihood of comprehending the statement metaphorically decreases, and therefore the likelihood of categorization also decreases. Thus, preference for the metaphor (categorical) form should be greater for highly apt statements than for less apt statements.

Method

Participants

For each experiment reported in this article, all of the participants were undergraduates at the University of Georgia, all received partial course credit for participation, and all were native speakers of English. Forty-eight undergraduates participated in the present experiment, and an additional 145 participated in the stimulus norming tasks described below.

Stimulus norming

A sample of 100 metaphors from Katz, Paivio, Marschark, and Clark (1988), McGlone (1996), and McGlone and Manfredi (2001) were subjected to several stimulus norming pre-tests. To manipulate aptness, we used the original 100 metaphors as the high apt stimuli (e.g., *A rooster is an alarm clock*), and for each metaphor we also created a low apt stimulus by substituting a less apt topic (e.g., *A robin is an alarm clock*). Thus there were 100 pairs of metaphors, with each pair consisting of one high apt and one low apt version, both of which had the same vehicle. We collected aptness ratings for all 200 metaphors. To manipulate conventionality, we had another group of participants generate a property that is com-

monly associated with each of the 100 vehicle concepts, and finally we had yet another group of participants rate how conventional those properties were for their given vehicle concept. These conventionality ratings were then used to distinguish between novel and conventional vehicles. Each of these stimulus norming procedures is described separately below.

Aptness ratings. Sixty-two undergraduates rated the aptness of each statement. Given the high positive correlation between aptness ratings of metaphors and of similes ($r = +.85$, Chiappe et al., 2003), all stimuli were presented in metaphor form. Two stimulus lists were created, so that the high apt and low apt versions of each pair always appeared on separate lists. Thus, each participant rated 50 high apt and 50 low apt metaphors on a scale that ranged from 1 (not at all apt) to 7 (highly apt). Following prior research, aptness was defined as "the extent to which the statement captures important features of the topic" (Chiappe et al., 2003, p. 97).

Conventional property generation. Fifty undergraduates participated in a property generation task, the aim of which was to identify the one property most commonly associated with each vehicle. Both metaphors of a pair were presented simultaneously onscreen (on separate lines), and each participant generated a single property that was conveyed by those metaphors. To emphasize that both metaphor topics should be considered, the following instructions were provided:

For this experiment, you will be presented with 80 pairs of metaphors for which you are to type in a SINGLE WORD that best captures the metaphorical meaning of BOTH metaphors. For instance, if you read the pair 'That race car was a rocket' and 'That bicycle was a rocket', you might type in 'fast' for the metaphorical property, as both bicycles and race cars can be fast.

Both metaphors remained on the screen as participants typed their responses into a textbox. For each pair of metaphors, we selected the most frequently pro-

vided response as the conventional property. These properties were subsequently used in a conventionality rating task.

Conventionality rating task. Following the procedure of Gentner and colleagues (e.g., Bowdle & Gentner, 2005; Gentner & Wolff, 1997), 33 undergraduates rated the conventionality of the most frequently generated property (see above) for each vehicle. We defined and exemplified conventionality as "the extent to which the concept is associated with a figurative meaning. For instance, BLUEPRINT has the well-established metaphoric interpretation of anything that provides a plan." Participants were asked to rate how conventional it is to use the concept (e.g., BLUEPRINT) to represent the given property (e.g., provides a plan). The scale ranged from 1 (very unconventional) to 7 (highly conventional).

Stimuli

From this extensive stimulus norming process we selected 64 pairs of high and low apt statements, 32 of which had a novel vehicle and 32 of which had a conventional vehicle. Stimuli were selected so as to be well controlled for aptness and conventionality within each condition, while maximally varying aptness and conventionality between conditions. Mean conventionality and aptness ratings are presented in Table 2. The conventionality ratings were submitted to an independent samples t test, which confirmed a significant effect [$t(62) = 12.10$, $p < .001$], thus indicating that the conventional vehicles ($M = 5.14$, $SE = .08$) were indeed more conventional than the novel vehicles ($M = 3.42$, $SE = .12$).

The aptness ratings were submitted to a 2 (Aptness) × 2 (Conventionality) mixed ANOVA, which confirmed a significant main effect of Aptness [$F(1, 62) = 294.71$, $p < .001$], such that the high apt items ($M = 4.85$, $SE = .12$) were indeed more apt than the low apt items ($M = 3.09$, $SE = .11$). Neither the main effect of Conventionality [$F(1, 62) = .18$, $p = .67$] nor its interaction with Aptness [$F(1, 62) = 1.62$, $p = .21$] was reliable. The lack of an interaction suggests that the manipulation of aptness did not differ appreciably between the novel stimuli and the conventional stimuli.

Table 2
Mean conventionality and aptness ratings

Condition	Conventionality	Aptness	Example
Conventional, high apt	5.14 (.08)	4.87 (.18)	Some runners are cheetahs.
Conventional, low apt	5.14 (.08)	2.98 (.17)	Some skaters are cheetahs.
Novel, high apt	3.42 (.12)	4.83 (.15)	That fashion model is a rail.
Novel, low apt	3.42 (.12)	3.20 (.15)	That football player is a rail.

Note: Standard errors are in parentheses. Conventionality Scale = 1 (very unconventional) to 7 (highly conventional). Aptness Scale = 1 (not at all apt) to 7 (highly apt).

Procedure

On each trial the metaphorical form and simile form of a statement appeared in white font on opposite sides of the computer display. The left-right location of metaphors and similes was counterbalanced across experimental lists. Below the two statements was a 7-point number line shown in red font. The number 1 appeared directly below the statement on the left of the display, and the number 7 appeared directly below the statement on the right of the display, with numbers 2 through 6 evenly spaced between. Either the high apt or the low apt version (but not both) of each statement appeared in each list. Hence, four lists were created to counterbalance the form presentation order and the aptness of each item. Each list included 16 stimuli from each of the four conditions defined by the 2 (Conventionality) × 2 (Aptness) design. Presentation of the 64 trials was randomized within-participants.

Participants indicated their form preference for each figurative statement by pressing the number (from 1 to 7) that corresponded with their preferred form. Specifically, they were instructed to indicate which form sounded more natural or sensible. They were further told that the stronger their preference for the statement on the left side, the closer their answer should be to 1, and the stronger their preference for the statement on the right side, the closer their answer should be to 7.

Results and discussion

Data were analyzed via separate 2 (Conventionality) × 2 (Aptness) ANOVAs with participants (F_p) and items (F_i) random. Both factors were within-participants and between-items. Categorical form (i.e., metaphor) preference was the dependent variable. Two of the four lists were reverse-scored so that higher numbers indicated greater categorical preference. Mean categorical preference ratings are shown in Fig. 1.
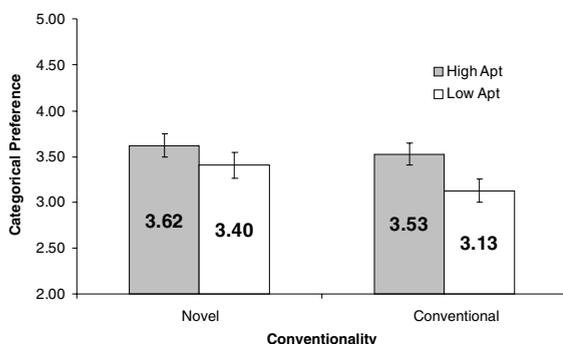

Fig. 1. Categorical preference ratings, Experiment 1. *Note.* Scale = 1 (Comparative form preferred) to 7 (Categorical form preferred). Error bars represent one standard error of the mean.

The career model predicted that categorical preference should be higher for conventional items than for novel items. As illustrated in Fig. 1, however, categorical preference was actually lower for the conventional items ($M = 3.33$, $SE = .11$) than the novel items ($M = 3.51$, $SE = .12$), [$F_p(1,47) = 12.61$, $p = .001$ and $F_i(1,124) = .97$, $p = .33$]. The present result is therefore opposite to that of Bowdle and Gentner (2005, Experiment 1), whose stimulus set confounded aptness with conventionality. In the present experiment conventionality and aptness were orthogonal, and in this case results failed to support the career model of metaphor.

The categorization model, on the other hand, predicted a main effect of Aptness. Consistent with this prediction, metaphor preference was greater for high apt items ($M = 3.57$, $SE = .12$) than for low apt items ($M = 3.27$, $SE = .13$), [$F_p(1,47) = 13.07$, $p = .001$ and $F_i(1,124) = 2.71$, $p = .10$]. That is, aptness rather than conventionality predicted metaphor preference. Results thus supported the categorization model, though the marginality of the item analysis renders this result less than conclusive. The interaction between Aptness and Conventionality was marginal by participants [$F_p(1,47) = 3.82$, $p = .06$] but was nonsignificant by items [$F_i(1,124) = 2.71$, $p = .62$].

Experiment 2

In Experiment 2 we measured comprehension latencies and collected ease-of-comprehension ratings. Following the methodology of Bowdle and Gentner (2005, Experiment 2), we used a 2 (Conventionality: conventional, novel) × 2 (Aptness: high, low) × 2 (Form: metaphor, simile) within-subjects design. As explained by Bowdle and Gentner (p. 202), conventional metaphors should be understood quickly and easily, because conventional vehicles are processed categorically and metaphors have a categorical form. Novel similes should also be understood quickly and easily, because novel vehicles induce comparison and similes have a comparative form. But for conventional similes and novel metaphors, their grammatical form is contrary to their presumed processing tendency. That is, a conventional simile has the form of a comparative statement, but its conventionality evokes categorical processing (and vice versa for novel metaphors). Thus, the career of metaphor predicts an interaction between Grammatical Form and Conventionality. Alternatively, the categorization model (Glucksberg & McGlone, 1999) predicts that high apt items should be comprehended faster and easier than low apt items, regardless of their conventionality and grammatical form (i.e., a main effect of Aptness).

Method

Participants

Sixty undergraduates participated.

Materials and design

The same 64 high- and low-apt item pairs from Experiment 1 were used in the current experiment. Four stimulus lists were created, so that each list contained either the high apt or the low apt version of each stimulus, and either the metaphor or the simile form of each stimulus. No concept appeared in any list more than once. Thus, each list included eight stimuli from each of the eight conditions defined by the 2 (Conventionality) × 2 (Aptness) × 2 (Form) design.

Because response times served as a dependent variable in this experiment, we conducted a 2 (Aptness) × 2 (Conventionality) × 2 (Form) ANOVA on the length (i.e., number of syllables) of the items. None of the interactions approached significance. The main effect of Aptness was also nonsignificant [$F(1, 248) = .89$, $p = .35$]. However, the main effect of Conventionality was significant [$F(1, 248) = 22.32$, $p < .001$]. The conventional items ($M = 7.02$, $SE = .14$) were shorter in length than the novel items ($M = 7.95$, $SE = .14$). Due to the additional word "like," the similes ($M = 7.98$, $SE = .14$) were longer in length than the metaphors ($M = 6.98$, $SE = .14$), [$F(1, 248) = 25.40$, $p < .001$]. We therefore treated length as a covariate where appropriate in the experiment proper.

Procedure

The procedure was an exact replication of that used by Bowdle and Gentner (2005, Experiment 2), with the sole exception that we additionally collected ease-of-interpretation ratings. On each trial, a row of asterisks (e.g., **** ** * *****) appeared for 500 ms and was then immediately replaced by the figurative statement. Participants were instructed to read the statement and press the spacebar once they had an interpretation in mind. Upon depression of the spacebar, a textbox appeared, and participants typed their interpretation. Participants then rated how easy it was to think of that interpretation on a scale from 1 (very difficult) to 7 (very easy). Presentation of the 64 trials was randomized within-participants. Eight practice trials preceded the experimental trials.

Results and discussion

Data were analyzed via 2 (Conventionality) × 2 (Aptness) × 2 (Grammatical Form) repeated measures ANOVAs. Comprehension times (i.e., from stimulus onset until depression of the spacebar) and ease-of-interpretation ratings served as dependent measures. Following the procedure of Bowdle and Gentner (2005), trials in which comprehension times were greater than 12 s and trials with missing interpretations were excluded from all analyses (6.8% of trials). Mean comprehension times and ease-of-interpretation ratings are shown in Figs. 2 and 3, respectively.

The 3-way interaction did not obtain for either the comprehension times [$F_p(1, 59) = .69$, $p = .41$ and $F_i(1, 124) = .58$, $p = .45$] or the interpretation ease ratings [$F_p(1, 59) = .06$, $p = .80$ and $F_i(1, 124) = .21$, $p = .65$]. Overall, metaphors ($M = 5.34$, $SE = .06$) and similes ($M = 5.28$, $SE = .07$) were equally easy to interpret [$F_p(1, 59) = 1.95$, $p = .17$ and $F_i(1, 124) = 1.95$, $p = .17$], but the metaphors ($M = 3643$, $SE = 98$) were
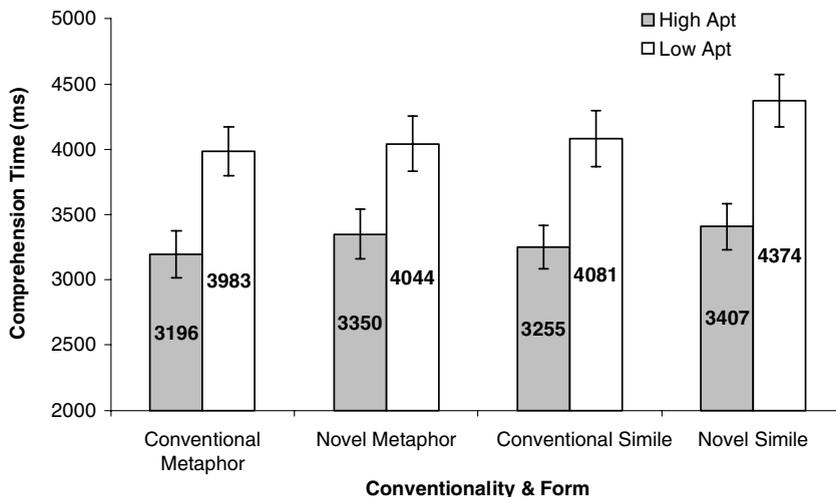

Fig. 2. Comprehension times, Experiment 2. *Note.* Error bars represent one standard error of the mean.
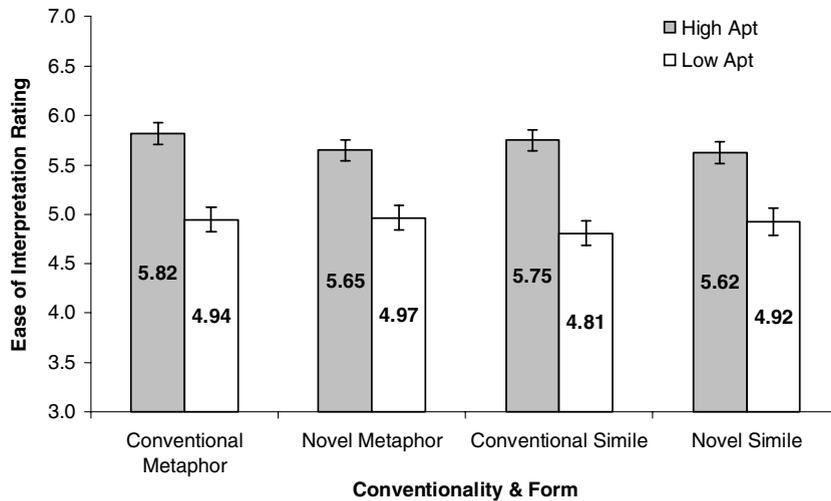
Fig. 3. Ease of interpretation ratings, Experiment 2. *Note*. Scale = 1 (very difficult) to (very easy). Error bars represent one standard error of the mean.

comprehended significantly faster than the similes ($M = 3779$, $SE = 99$), [$F_p(1, 59) = 4.66$, $p < .05$ and $F_i(1, 124) = 5.50$, $p < .05$]. The longer time necessary to comprehend the similes may be attributable to the additional word "like." In support of this explanation, length reliably predicted mean comprehension times, $r = +.36$, $p < .001$. Alternatively, the comparative processing that was presumably induced by the similes may be more time consuming than the categorical processing of the metaphors. To test these possibilities, we conducted an ANCOVA with length as the covariate and Grammatical Form as a between-items factor (collapsing across aptness). The main effect of Form was eliminated, $F_i(1, 125) < .01$, $p = .95$, indicating that the additional word "like" was responsible for the similes' longer comprehension times. Other interaction and main effect analyses are reported below according to their theoretical relevance.

Conventionality

Conventionality did not mediate metaphorical processing. Recall that the career model of metaphor predicts an interaction of Conventionality and Grammatical Form, such that novel similes should be comprehended easier than novel metaphors, whereas conventional metaphors should be comprehended easier than conventional similes (Bowdle & Gentner, 2005, p. 202). This prediction was unsupported. First, the interaction between Conventionality and Grammatical Form was nonsignificant in both the comprehension time [$F_p(1, 59) = .76$, $p = .39$ and $F_i(1, 124) = .52$, $p = .47$] and ease-of-interpretation analyses [$F_p(1, 59) = .56$, $p = .46$ and $F_i(1, 124) = .33$, $p = .57$]. Second, planned comparisons revealed that novel similes ($M = 3891$, $SE = 173$) were actually com-

prehended more *slowly* than novel metaphors ($M = 3697$, $SE = 189$), $F_p(1, 59) = 3.67$, $p = .06$ and $F_i(1, 62) = 4.19$, $p < .05$. Third, comprehension times were approximately equivalent for the conventional metaphors ($M = 3590$, $SE = 169$) and the conventional similes ($M = 3668$, $SE = 176$), $F_p(1, 59) = .94$, $p = .34$ and $F_i(1, 62) = 1.50$, $p = .23$. The planned comparisons with ease ratings were nonsignificant for both the novel statements [$F_p(1, 59) = .20$, $p = .65$ and $F_i(1, 62) = .46$, $p = .50$] and the conventional statements [$F_p(1, 59) = 2.51$, $p = .12$ and $F_i(1, 62) = 1.55$, $p = .22$].

The main effect of Conventionality was not significant in the ease ratings [$F_p(1, 59) = .71$, $p = .40$ and $F_i(1, 124) = .30$, $p = .58$], and in the comprehension times it was significant only in the participant analysis [$F_p(1, 59) = 5.57$, $p < .05$ and $F_i(1, 124) = 1.58$, $p = .21$]. Given the importance of conventionality for the career model of metaphor, we examined this effect in the participant analysis more closely. Conventional similes were comprehended faster than novel similes [$F_p(1, 59) = 4.61$, $p < .05$], but conventional and novel metaphors were comprehended equally fast [$F_p(1, 59) = 1.48$, $p = .23$]. Thus, the effect of Conventionality held only for the similes, only in comprehension times, and only in the participant analysis. Essentially, there was no evidence that conventionality predicts metaphor comprehension. So we again failed to replicate the result of Bowdle and Gentner (2005). This finding suggests that their result was more likely attributable to aptness than to conventionality.

Aptness

Aptness facilitated both the speed and the ease of comprehension. The main effect of Aptness was highly

significant in the analyses of both comprehension times [$F_p(1, 59) = 77.05$, $p < .001$ and $F_i(1, 124) = 34.88$, $p < .001$] and ease-of-interpretation ratings [$F_p(1, 59) = 121.14$, $p < .001$ and $F_i(1, 124) = 57.69$, $p < .001$]. Specifically, the high apt statements ($M = 3302$, $SE = 88$) were comprehended 819 ms faster than the low apt statements ($M = 4121$, $SE = 101$). These high apt statements ($M = 5.71$, $SE = .05$) were also rated easier to interpret than the low apt statements ($M = 4.91$, $SE = .06$). Furthermore, the facilitative effect of aptness on comprehension held regardless of whether the statement appeared in metaphor or simile form, as the interaction between Aptness and Grammatical Form failed to approach significance for either comprehension times [$F_p(1, 59) = 1.44$, $p = .24$ and $F_i(1, 124) = 1.65$, $p = .20$] or ease-of-interpretation ratings [$F_p(1, 59) = .14$, $p = .71$ and $F_i(1, 124) = .15$, $p = .70$]. Finally, as clearly illustrated in Figs. 2 and 3, the facilitative effect of aptness held regardless of whether the statement was novel or conventional. The interaction between Aptness and Conventionality failed to approach significance for the comprehensions times [$F_p(1, 59) = .06$, $p = .81$ and $F_i(1, 124) = .09$, $p = .77$] and was reliable only in the participant analysis for the ease-of-interpretation ratings [$F_p(1, 59) = 4.34$, $p = .04$ and $F_i(1, 124) = .81$, $p = .37$]. Thus, results indicated that aptness mediated metaphor comprehension. The categorization model was therefore supported.

Experiment 3

In Experiment 3 we used the metaphorical categorization task (Jones & Estes, 2005), whereby participants judge whether the metaphor topic is a member of the category named by the vehicle. For example, after reading *Some runners are cheetahs*, participants rated the extent to which RUNNERS belong in the category of CHEETAHS. Although we previously found that highly apt metaphors were more likely to induce category membership (Jones & Estes, 2005, Experiment 3), conventionality was neither tested nor controlled in that study. Thus, the present experiment had a 2 (Conventionality) x 2 (Aptness) within-participants design. The career model predicts that category membership ratings should be higher following conventional metaphors, since novel metaphors cannot be processed categorically (Bowdle & Gentner, 2005). The categorization model, in contrast, predicts that category membership should be greater following highly apt metaphors, since less apt metaphors are simply less likely to be comprehended metaphorically (Glucksberg, 2003). In addition to the experimental stimuli, we also included scrambled metaphors (e.g., *That baby is a*

map) and borderline category statements (e.g., *Poker is a sport*) as a validity check.

Method

Participants

Thirty-one undergraduates participated.

Materials and design

The metaphor form of the 128 items used in the previous two experiments comprised the experimental items. Two lists were constructed so that each participant saw either the low apt or high apt version of each metaphor. Each list included 16 stimuli from each of the four conditions defined by the 2 (Conventionality) × 2 (Aptness) design. We also included 32 scrambled metaphors (e.g., *Hard work is a teddy bear*) and 32 borderline literal items (e.g., *A jet ski is a boat*) used in our previous study (Jones & Estes, 2005, Experiment 3) to serve as a manipulation check. Category membership should be the lowest for the scrambled metaphors, because there are no salient properties shared between topic (e.g., HARD WORK) and vehicle (e.g., TEDDY BEAR). The literal statements included exemplars (e.g., JET SKI) that have properties somewhat typical of their categories (e.g., BOAT), and therefore were predicted to have category membership ratings higher than those for the experimental metaphors.

Procedure

Each stimulus (e.g., *A robin is an alarm clock*) was displayed for 2 s in bold red 22-point font on a black background. Participants indicated the extent to which the topic concept was a member of the vehicle-named category (e.g., To what extent is a ROBIN a member of the category ALARM CLOCK?) on a scale from 1 (not at all a member) to 7 (completely a member). The order of trials was randomized within-participants.

Results and discussion

Relative to the experimental items ($M = 3.01$, $SE = .10$), the scrambled metaphors ($M = 1.62$, $SE = .17$) induced lower category membership ratings and the borderline literal items ($M = 5.17$, $SE = .12$) induced higher category ratings.[3] This suggests that the metaphorical categorization paradigm provided a valid measure of category membership (see Jones & Estes,

---

[3] Category membership ratings were significantly lower for the scrambled metaphors ($M = 1.62$, $SE = .17$) than for both the low apt conventional metaphors [$M = 2.32$, $SE = .18$, $t_p(30) = 7.68$, $p < .001$ and $t_i(94) = 5.26$, $p < .001$] and the low apt novel metaphors [$M = 2.45$, $SE = .21$, $t_p(30) = 8.27$, $p < .001$ and $t_i(94) = 6.66$, $p < .001$]. Thus, the low apt items were not anomalous.
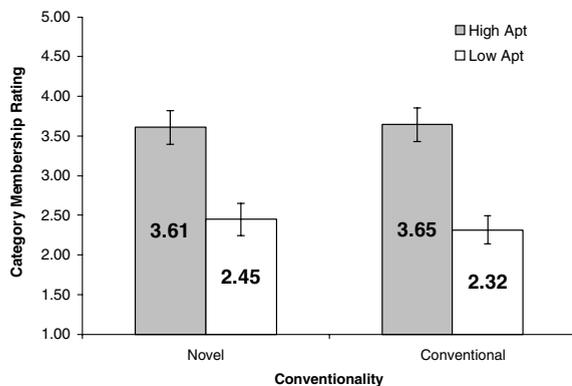
Fig. 4. Category membership ratings, Experiment 3. *Note.* Scale = 1 (not at all a member) to 7 (completely a member). Error bars represent one standard error of the mean.

2005). Data for the experimental items were therefore analyzed via 2 (Conventionality) × 2 (Aptness) repeated measures ANOVAs. Mean category membership ratings are illustrated in Fig. 4.

As shown in the figure, category ratings were equivalent between the conventional ($M = 2.98$, $SE = .19$) and novel metaphors ($M = 3.03$, $SE = .20$), $F_p(1, 30) = .45$, $p = .51$ and $F_i(1, 124) = .08$, $p = .78$. That is, conventionality exerted no effect on categorical processing, and hence the career of metaphor was not supported (see also Jones & Estes, 2005). As predicted by the categorization model, however, category membership ratings were higher for the high apt ($M = 3.63$, $SE = .20$) than for the low apt metaphors ($M = 2.38$, $SE = .19$), $F_p(1, 30) = 133.69$, $p < .001$ and $F_i(1, 124) = 55.86$, $p < .001$. Furthermore, this effect of Aptness was consistent across the novel and conventional metaphors, as evidenced by the lack of an interaction between Conventionality and Aptness, $F_p(1, 30) = 1.35$, $p = .26$ and $F_i(1, 124) = .25$, $p = .62$. Clearly, aptness rather than conventionality predicted categorical processing. This result is thus consonant with the previous two experiments in supporting the categorization model over the career model.

General discussion

Empirical summary

Across three experiments, conventionality exerted no reliable effect on metaphorical processing. Relative to novel statements, conventional statements were *not* more likely to be preferred in categorical form (see Fig. 1), were *not* comprehended faster (see Fig. 2), were *not* comprehended more easily (see Fig. 3), and were *not* more likely to induce categorization (see Fig. 4). On the other hand, aptness did consistently affect metaphorical

processing. Relative to less apt statements, highly apt statements were more likely to be preferred in categorical form (Fig. 1), were comprehended faster (Fig. 2), were comprehended more easily (Fig. 3), and were more likely to induce categorization (Fig. 4). These effects were highly significant and in general quite large. Evidently, the effects of conventionality reported by Bowdle and Gentner (2005) may instead be attributable to aptness, since those two factors were confounded in their stimulus set. In the present experiments we manipulated conventionality and aptness independently, and only aptness exhibited any discernible effect on metaphorical processing.

To more directly compare our results to the prior studies summarized in Table 1, we calculated bivariate correlations between the aptness ratings collected in Experiment 1 and each of the four dependent measures of Experiments 1 through 3. In Experiment 1, aptness reliably predicted metaphor preference [$r = +.39$, $p < .001$], such that more apt statements were more likely to be preferred in metaphorical form. This significant positive correlation corroborates several previous investigations (see Table 1). In Experiment 2, aptness reliably predicted comprehension latencies [$r = -.60$, $p < .001$], with the negative correlation indicating that more apt statements were comprehended faster than less apt statements. This result is again consistent with some prior studies. Aptness was also significantly correlated with ease-of-interpretation ratings [$r = +.65$, $p < .001$] in Experiment 2. That is, as aptness increased, so did comprehensibility. Several previous researchers have reported similar results (see Table 1). Finally, in Experiment 3, aptness was positively related to category membership [$r = +.65$, $p < .001$]. The more apt the metaphor, the more likely it was to induce categorization. This finding replicates our previous result (Jones & Estes, 2005), but unlike that prior study, the present effect of aptness cannot be attributed to conventionality.

Theoretical evaluation

These results pose a serious challenge to the career of metaphor. The manipulations of conventionality and aptness were equally strong (see Table 2), yet only aptness reliably predicted metaphor comprehension. Of course, the failure of conventionality to mediate comprehension in these experiments does not imply that conventionality has no influence on metaphor comprehension. But if conventionality does affect comprehension, then either it must require a stronger manipulation than aptness, or its influence must be subtle compared to that of aptness. In either case, these results suggest that conventionality is not the primary predictor of metaphor comprehension, and therefore the career model was not supported.

Results instead support the categorization model, which claims that aptness is the primary predictor of comprehension. This model posits an interaction of the topic and vehicle concepts (cf. Black, 1962, 1993), such that the topic guides the selection of candidate properties for attribution from the vehicle (Glucksberg et al., 1997; McGlone & Manfredi, 2001). Because aptness takes into account both the figurative property of the vehicle and the dimensions of the topic, it follows that aptness should predict metaphor comprehension (Glucksberg & Keysar, 1990, pp. 13–14; Glucksberg & McGlone, 1999, p. 1544; Glucksberg et al., 1997, pp. 57–59). The present experiments clearly supported this prediction.

To be clear, the career model also posits that metaphor comprehension is a directional process, with properties being projected from vehicle to topic. However, the career model is not interactive, in the sense that the model specifies no active role for the topic during property projection. Indeed, the career model does not attribute any explanatory power to aptness, and in fact Bowdle and Gentner (2005, pp. 204–205) explicitly sought to reject aptness as an explanation of their results. Instead, the career model claims that comprehension is mediated by conventionality, which is determined by the vehicle concept alone. Thus the primary flaw of the career model is its failure to capture adequately the different roles of the topic and the vehicle concepts. Only the categorization model, then, predicts that aptness should facilitate metaphor comprehension, and that conventionality should not.

Although these critical predictions of the categorization model were supported, it must be noted that these experiments provide only limited evidence that the metaphors were processed *categorically*. In Experiment 1, for instance, preference for the categorical form (i.e., metaphor) was indeed greater for highly apt statements than for less apt statements, but even those highly apt statements were below the scalar midpoint of 4 (see Fig. 1). This suggests that both groups of statements were actually preferred in the comparative form (i.e., simile). In fact, this finding of a general preference for similes over metaphors has been obtained consistently across several studies (e.g., Bowdle & Gentner, 1999, 2005; Chiappe & Kennedy, 1999; Chiappe et al., 2003), and it remains an outstanding issue for proponents of the categorization model to explain. In Experiment 2, highly apt metaphors were indeed comprehended faster and easier than less apt metaphors, but this only can be taken as evidence of categorical processing if one assumes that category statements are comprehended faster and easier than comparison statements. Finally, in Experiment 3 category membership ratings were indeed higher for apt metaphors than for unapt metaphors, but again the ratings were below the scalar midpoint (see Fig. 4). Thus, the present experiments do not provide strong evidence of categorical processing per se.

These results raise the interesting theoretical question of whether the career of metaphor can be preserved by substituting aptness for conventionality in its predictions. The primary tenet of the career of metaphor is that the means of comprehension shifts from comparative processing to categorical processing. If this shift in processing were predicated on aptness rather than conventionality, would the model be tenable? What would be required is a demonstration that highly apt and less apt metaphors were processed via distinct processes of categorization and comparison, respectively. But the present experiments do not indicate that apt and unapt metaphors are processed in qualitatively different ways; these results indicate only quantitative differences in the processing of apt and unapt metaphors. So unfortunately, there is currently no more reason to accept the new career model than the old career model. Nevertheless, this possibility seems a particularly important avenue for more direct investigation.

Methodological implication

Importantly, the present results indicate that aptness and conventionality are independent factors with differential contributions to metaphor comprehension. As outlined in the introduction, the great majority of studies in the literature have manipulated one of these factors without controlling the other. This has culminated in the contrast of two opposing models, one with aptness as the primary mediator of metaphor comprehension (Glucksberg & McGlone, 1999) and the other with conventionality as the hypothesized mediator (Bowdle & Gentner, 2005). But because the previous tests of these models have routinely confounded the two factors, it has been impossible to discriminate between the models. The present results therefore indicate that future investigations of metaphor comprehension should either manipulate or control both aptness and conventionality. We hope that the utilization of better-controlled stimulus sets will foster the advancement of more precise models of metaphor comprehension. Indeed, we hope that the stimulus set included in the Appendix may contribute to this pursuit.

The generality of aptness

In addition to its role in nominal metaphor, aptness also appears to mediate the comprehension of nominal combinations such as RAINBOW TROUT and MUSHROOM CLOUD. Indeed, many researchers suggest that metaphor

and conceptual combination involve the same basic processes of comprehension (Estes & Glucksberg, 2000; Gagné, 2002; Gagné, Friedman, & Faries, 1996; Wisniewski, 1997). Just as metaphor comprehension entails attributing a salient property of the vehicle to a relevant dimension of the topic, conceptual combination involves the attribution of a salient property of the modifier concept to a relevant dimension of the head concept. Estes and Glucksberg (2000) showed that for combinations with highly relevant head noun dimensions (e.g., weight for LUGGAGE) and highly salient modifier properties (e.g., light for FEATHER), property interpretations (i.e., "luggage that is light") were provided 79% of the time. But when either the property of the modifier was not salient (e.g., COTTON LUGGAGE, 16%) or the dimension of the head noun was not relevant (e.g., FEATHER STORAGE, 23%), then metaphorical property interpretations were much less likely (see also Bock & Clifton, 2000; Katz, 1982). Thus, aptness appears to mediate comprehension of several linguistic tropes including metaphor, simile, and conceptual combination.

Appendix

Experimental stimuli

Conv. Rating	High Apt Metaphor	Aptness Rating	Low Apt Metaphor	Aptness Rating
Novel				
1.79	A business is a living organism.	3.76	A kitchen is a living organism.	2.85
1.82	A goalie is a spider.	3.27	A fisherman is a spider.	1.76
2.42	His old car is a boat.	3.21	His old lawn mower is a boat.	2.61
2.58	Beavers are lumberjacks.	5.24	Termites are lumberjacks.	4.18
2.64	Some stomachs are barrels.	4.70	Some bladders are barrels.	3.59
2.67	Some lectures are sleeping pills.	6.09	Some comedians are sleeping pills.	3.31
2.97	Research is mountain climbing.	4.27	Grading is mountain climbing.	2.90
3.00	Her unflinching gaze is ice.	4.48	Her quick nod is ice.	3.61
3.00	Her mother is a mule.	3.17	Her hairdresser is a mule.	2.03
3.18	Insults are razors.	5.41	Evaluations are razors.	3.52
3.18	Smoking is a time bomb.	5.24	Cocaine is a time bomb.	4.66
3.21	Education is a lantern.	5.17	Debate is a lantern.	3.18
3.45	Her husband is a gem.	4.97	Her ex-husband is a gem.	3.64
3.48	A rooster is an alarm clock.	5.52	A robin is an alarm clock.	3.88
3.52	My computer skills course is a joke.	5.48	My organic chemistry course is a joke.	4.31
3.55	The mind is an arena.	4.72	The subconscious is an arena.	4.03
3.55	Many stores are jungles.	4.36	Many beaches are jungles.	2.24
3.58	His kindergarten class is a zoo.	5.69	His college class is a zoo.	4.64
3.64	The driveway is an ice rink.	4.00	The front lawn is an ice rink.	2.86
3.73	Her final exam in Geography was a bear.	3.41	Having summers off was a bear.	2.12
3.73	A lie is a dagger.	5.69	A wish is a dagger.	2.48
3.85	The bad news was an earthquake.	4.79	The good news was an earthquake.	2.00
3.94	A zoo is a museum.	4.14	A conference is a museum.	2.48
3.97	Adventure is a roller coaster.	5.66	Traveling is a roller coaster.	4.79
4.03	That criminal's fingerprint is a portrait.	4.85	That criminal's pathway is a portrait.	2.79
4.09	Some teachers are encyclopedias.	5.64	Some coaches are encyclopedias.	3.31
4.09	My grandmother is a peach.	4.82	My boyfriend is a peach.	3.41
4.12	Jalapeño peppers are fire.	5.03	Bell peppers are fire.	3.97
4.12	Some divorces are storms.	5.70	Some adoptions are storms.	2.59
4.15	Sarcasm is a veil.	4.67	Hostility is a veil.	3.79
4.18	Music can be medicine.	6.18	Reality TV can be medicine.	2.69
4.21	That fashion model is a rail.	5.30	That football player is a rail.	2.28
4.36	Alcohol is a crutch.	4.52	Pizza is a crutch.	2.33
4.39	Dancers can be butterflies.	4.66	Soccer players can be butterflies.	2.64
4.61	Discipline is fertilizer.	3.38	Control is fertilizer	2.73
4.61	A tree is an umbrella.	4.21	A magazine is an umbrella.	2.15
4.64	Love is a flower.	5.34	Hatred is a flower.	1.85
4.64	Some voices are sirens.	5.00	Some snores are sirens.	3.76

Appendix (*continued*)

Conv. Rating	High Apt Metaphor	Aptness Rating	Low Apt Metaphor	Aptness Rating
4.70	That receptionist is a breath of fresh air.	4.52	That accountant is a breath of fresh air.	3.38
4.79	A best friend is an anchor.	6.07	An opponent is an anchor.	2.94
4.79	His marriage was a short leash.	4.00	His office was a short leash.	2.52
4.79	The cheering crowd was thunder.	5.39	The basketball player was thunder.	2.76
4.82	The Great Plains are a board.	2.76	The coastal beaches are a board.	1.90
4.94	A college degree is a doorway.	6.07	A course completion is a doorway.	4.24
4.94	Some smiles are magnets.	5.39	Some tears are magnets.	3.00
5.03	My young cousin is a shrimp.	4.85	My father is a shrimp.	3.52
5.06	Butlers can be donkeys.	2.64	Grandparents can be donkeys.	1.79
5.06	Some fashion models are twigs.	6.15	Some swimmers are twigs.	3.24
5.09	Love is a journey.	6.48	A dream is a journey.	5.69
5.21	That pregnant woman is a duck.	2.82	That professor is a duck.	2.00
5.27	Some dogs are princesses.	4.33	Some hamsters are princesses.	1.97
5.30	My cat's fur is silk.	4.88	My rat's fur is silk.	2.76
5.30	Books are treasure chests.	5.66	Tabloids are treasure chests.	2.67
5.45	Some runners are cheetahs.	5.64	Some skaters are cheetahs.	2.62
5.48	Many jobs are jails.	4.88	Many teams are jails.	1.93
5.52	Time is money.	5.88	Knowledge is money.	5.00
5.55	The planet Earth is a ball.	4.61	The nearest star is a ball.	2.45
5.61	Birds are airplanes.	3.70	Bats are airplanes.	2.14
5.73	Ideas can be diamonds.	5.62	Reviews can be diamonds.	3.91
5.73	Anger is a volcano.	5.62	Sadness is a volcano.	3.76
5.76	The senator is a fossil.	4.52	The track star is a fossil.	2.24
5.76	My boyfriend's arms are steel.	5.36	My grandfather's legs are steel.	3.45
5.82	That bedroom is a dump.	5.52	That conference room is a dump.	4.79
5.85	Memory is a warehouse.	5.52	Intelligence is a warehouse.	3.38

References

Aisenmann, R. A. (1999). Structure-mapping and the simile-metaphor preference. *Metaphor and Symbol, 14*, 45–51.

Black, M. (1962). Metaphor. In M. Black (Ed.), *Models and metaphors* (pp. 25–47). Ithaca, New York: Cornell University Press.

Black, M. (1993). More about metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 19–41). Cambridge: Cambridge University Press.

Blank, G. D. (1988). Metaphors in the lexicon. *Metaphor and Symbolic Activity, 3*, 21–36.

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 295–308.

Bock, J. S., & Clifton, C. (2000). The role of salience in conceptual combination. *Memory & Cognition, 28*, 1378–1386.

Bowdle, B.F., Gentner, D. (1999). Metaphor comprehension: From comparison to categorization. In: Hahn, M., Stoness, S.C. (Eds.), Proceedings of the twenty-first annual conference of the cognitive science society (pp. 90–95). Mahwah, NJ: LEA.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review, 112*, 193–216.

Camac, M. K., & Glucksberg, S. (1984). Metaphors do not use associations between concepts, they are used to create them. *Journal of Psycholinguistic Research, 13*, 443–455.

Chiappe, D. L., & Kennedy, J. M. (1999). Aptness predicts preference for metaphors or similes, as well as recall bias. *Psychonomic Bulletin & Review, 6*, 668–676.

Chiappe, D. L., & Kennedy, J. M. (2001). Literal bases for metaphor and simile. *Metaphor and Symbol, 16*, 249–276.

Chiappe, D. L., Kennedy, J. M., & Chiappe, P. (2003). Aptness is more important than comprehensibility in preference for metaphors and similes. *Poetics, 31*, 51–68.

Chiappe, D. L., Kennedy, J. M., & Smykowski, T. (2003). Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol, 18*, 85–105.

Estes, Z., & Glucksberg, S. (2000). Interactive property attribution in concept combination. *Memory & Cognition, 28*, 28–34.

Gagné, C., Friedman, A., & Faries, J. (1996). Effect of priming on the comprehension of predicative metaphors. *Metaphor and Symbolic Activity, 11*, 125–143.

Gagné, C. L. (2002). Metaphoric interpretations of comparison-based combinations. *Metaphor and Symbol, 17*, 161–178.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D., & Bowdle, B. F. (2001). Convention, form, and figurative language processing. *Metaphor and Symbol, 16*, 223–247.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like Analogy. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 199–253). Cambridge, MA: MIT Press.

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language, 37*, 331–355.

Gerrig, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 9*, 667–675.

Gibb, H., & Wales, R. (1990). Metaphor or simile: Psychological determinants of the differential use of each sentence form. *Metaphor and Symbolic Activity, 5*, 199–213.

Gibbs, R. W. (1992). Categorization and metaphor understanding. *Psychological Review, 99*, 572–577.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8*, 183–206.

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Science, 7*, 92–96.

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

Glucksberg, S., & McGlone, M. S. (1999). When love is not a journey: What metaphors mean. *Journal of Pragmatics, 31*, 1541–1558.

Glucksberg, S., McGlone, M. S., & Manfredi, D. (1997). Property attribution in metaphor comprehension. *Journal of Memory and Language, 36*, 50–67.

Jones, L. L., & Estes, Z. (2005). Metaphor comprehension as attributive categorization. *Journal of Memory and Language, 53*, 110–124.

Katz, A. N. (1982). Metaphoric relationships: The role of feature saliency. *Journal of Psycholinguistic Research, 11*, 283–296.

Katz, A. N., Paivio, A., Marschark, M., & Clark, J. M. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity, 3*, 191–214.

Kusumi, T. (1987). Effects of categorical dissimilarity and affective similarity between constituent words on metaphor appreciation. *Journal of Psycholinguistic Research, 16*, 577–595.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

McGlone, M. S. (1996). Conceptual metaphors and figurative language interpretation: Food for thought? *Journal of Memory and Language, 35*, 544–565.

McGlone, M. S., & Manfredi, D. A. (2001). Topic-vehicle interaction in metaphor comprehension. *Memory & Cognition, 29*, 1209–1219.

Miller, G. A. (1993). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought* (pp. 357–400). Cambridge: Cambridge University Press.

Nakamoto, K., & Kusumi, T., (2004). The effect of repeated presentation and aptness of figurative comparisons on preference for metaphor forms. Proceedings of the twenty-sixth annual conference of the cognitive science society. Chicago, IL.

Sternberg, R. J., & Nigro, G. (1983). Interaction and analogy in the comprehension and appreciation of metaphors. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 35A*, 17–38.

Tourangeau, R., & Rips, L. (1991). Interpreting and evaluating metaphors. *Journal of Memory and Language, 30*, 452–472.

Tourangeau, R., & Sternberg, R. J. (1981). Aptness in metaphor. *Cognitive Psychology, 13*, 27–55.

Tourangeau, R., & Sternberg, R. J. (1982). Understanding and appreciating metaphors. *Cognition, 11*, 203–244.

Utsumi, A., & Kuwabara, Y. (2005). Interpretive diversity as a source of metaphor-simile distinction. In: Proceedings of the twenty-seventh annual conference of the cognitive science society. Stressa, Italy.

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review, 4*, 167–183.

Wolff, P., & Gentner, D. (2000). Evidence for role-neutral initial processing of metaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 529–541.

Zharikov, S., & Gentner, D. (2002). Why do metaphors seem deeper than similes. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the twenty-fourth annual conference of the cognitive science society* (pp. 976–981). Fairfax VA: George Mason University.