Joseph Glicksohn and Chanita Goodblatt

Empirical Studies of Poetic Metaphor

Abstract: This chapter has two primary goals: first, to present a concise survey of scholarship on empirical studies of poetic metaphor; and second, to define several major issues seen in such empirical studies that can guide future research in this field. In the survey, we distinguish among three types of experimental methodology employed in this interdisciplinary domain: reaction-time (RT) studies investigating, for example, semantic decisions; the use of rating scales assessing, for example, metaphor aptness; and the use of thinking-aloud protocols investigating, for example, the process of metaphor comprehension or the experience of affect. This review considers such questions as: Should poetic metaphor be studied in isolation from the context of the poetic text? Should similes and metaphors be treated equivalently in an empirical study? In the discussion of the methodologies, major questions raised include: Who should be the readers of the poetic text? Should the studies of poetic metaphor be necessarily informed both by literary scholarship and cognitive psychology? What are the methodological strengths and limitations of the empirical studies?

Introduction

From the outset, it should be noted that the authors of this chapter have a particular Gestalt-oriented approach to the study of poetic metaphor (Glicksohn & Goodblatt, 1993). Poetic metaphor will be defined for present purposes as a metaphor deriving from a poetic text. This Gestalt-oriented approach requires embedding the poetic metaphor within the poetic text, which provides the relevant context for its understanding. A poetic metaphor is therefore conceived as a *gestalt* appearing within a poetic text (Goodblatt & Glicksohn, 2010) and will have to be comprehended as part of *that* text – which itself is a *gestalt* (Glicksohn & Goodblatt, 1993). Wertheimer (2010, p. 50) depicts the structure of a *gestalt*:

> An integrated, articulated whole, a Gestalt is an organized totality within which the nature, place, role and function of each part is precisely what it must be, given the nature of the whole. The parts are in dynamic interaction with each other and with the whole; they are not a mere bundle or concatenation of items that happen to be arbitrarily glued or hooked together by accidental proximity in space and time.

https://doi.org/10.1515/9783110645958-006

Following Wertheimer, the poetic metaphor, which comprises a part of the greater whole of the poetic text, is in dynamic interaction both between its two parts as well as with the surrounding poetic text. Gräbe (1984, p. 434) clarifies this, writing: "a consideration of global aspects in a textual analysis of metaphorical expressions is a prerequisite for an adequate understanding of the complex interactional processes that typically characterize poetic metaphor (i. e. metaphorical expressions within poetic texts)." Snævarr (2010, p. 149) supports our position, writing that "Glicksohn and Goodblatt have compelling arguments in favor of metaphors being Gestalten. In my view, the same holds for poems (at least typical ones). Just like a Gestalt, a poem is a whole bigger than the sum of its parts."

In line with our theoretical orientation, we argue that a poetic metaphor is not the same as a poetic simile: A metaphor goes beyond the similarity of its *tenor*, *target*, or *primary subject* and its *vehicle*, *source*, or *secondary subject* (Black, 1962; Richards, 1936); rather it is created by an *interaction* between its primary and secondary subjects (Glicksohn & Goodblatt, 1993). Indeed, Kintsch (2000, p. 264) has suggested that "metaphors do not transfer a single feature, or even a small set of features, but rearrange a whole semantic field…the meaning of a metaphor involves a restructuring of the semantic space, which is more difficult to capture than simple feature transfer." Viewing metaphor as analogy or as implicit simile (Gentner et al., 2001) will presuppose the similarity of its topic and vehicle, and hence will rely on a process of comparison of topic and vehicle in metaphor comprehension (Glicksohn & Goodblatt, 1993; Glucksberg & Keysar, 1990; Holyoak & Stamenković, 2018). If, however, metaphor in general, and poetic metaphor in particular, is viewed as *generating* the similarity of its topic and vehicle, rather than relying on preexisting similarity, then the reader's task becomes one of *problem solving* (Goodblatt & Glicksohn, 2002).

These statements raise three major questions. First, why should poetic metaphor be read within its text? Franklin (1994, p. 263) insightfully writes that a Gestalt-oriented approach views *the poet* as drawing "deeply on the *expressive properties of language* – properties of sound and rhythm that directly convey feelings such as happiness or sadness. The poet draws also on the *structural possibilities* of syntactic arrangement and rearrangement." In other words, these aspects of the poetic text comprise a contextual *foregrounding* (Aryani et al., 2016; Gambino & Pulvirenti, 2019; Miall & Kuiken, 1994; Zyngier et al., 2007). This foregrounding will provide the type of structural equivalents (or constraints) that will help to induce a coherent reading of poetic metaphor. Van Peer et al., (2007, p. 211) have noted, however, that one must distinguish "between foregrounded elements … and foregrounding effects (the effects that foregrounded elements may have on some readers, for instance aesthetic pleas-

ure)." The richness and complexity of such a coherent reading of poetic metaphor stands in contrast to the reading of poetic metaphor (indeed any metaphor) that is presented in isolation.[1] As Jacobs and Kinder (2018, p. 94) have insightfully concluded from their own study, "it is likely that within the dynamic context of, say, the reading act of a Shakespeare sonnet a given metaphoric construction is processed differently than when presented in isolation."

Foregrounding, of course, also implies a background – the familiar figure-ground emphasis of *Gestaltpsychologie*. Following Jacobs (2015b), one can argue that if foregrounding elicits an *aesthetic* literary experience, then this appears on a background eliciting an *immersive* literary experience. This type of distinction has been discussed by Glicksohn and Berkovich-Ohana (2012). Poetic metaphor appearing in context can thus be viewed as eliciting an *analytic* literary experience (Goodblatt & Glicksohn, 2002, for example, specifically refer to *problem solving* here), coming on the background of an aesthetic literary experience, itself coming on the background of an immersive literary experience. One has to consider all of these levels in unison – not in isolation, rather in interaction.

The second question is: How much of the poetic text is required as context for the comprehension of a poetic metaphor? Empirical investigators who opt for a short textual context, might focus on studying brain activity, using metaphors as the basis for their task. Thus tasks required of the participants include: to "judge whether the presented two-word expression conveys a meaning (be it literal or metaphoric) or does not convey a meaning as a pair" (Arzouan et al., 2007, p. 224); to decide "whether the two words are related to each other or not" (Kennett et al., 2018, p. 4); or to judge "whether the word appearing on the screen seemed familiar or unfamiliar" (Forgács et al., 2012, p. 1435). Note that in these studies, the metaphor can appear *within* text but also *without* text (e. g., Mashal et al., 2005; Mashal et al., 2009; McCabe, 1983). In contradistinction to such an approach, our stand is that to study *metaphor comprehension*, the metaphor should be presented within the complete poetic text. In the recent study reported by Bambini et al. (2019), two-word poetic metaphors in genitive form (e. g., "theme of velvet") were presented within their poetic text. The task was

1 The study of poetic metaphor (also known as *literary* metaphor) appearing in isolation has a long history, a landmark study being that of Katz et al. (1985; see also Katz et al., 1988). Studies using such material include Glicksohn (1994) and Stamenković et al. (2019). Poetic metaphor presented in isolation might very well elicit a different mode of processing – at both the cognitive and electrophysiological levels – than that elicited by poetic metaphors presented in context. Consider the metaphor "My surgeon is a butcher," discussed by such senior theorists of metaphor as Gentner et al. (2001), Glucksberg and Keysar (1990), Kintsch (2000), and Kittay (1997). Without a relevant context, one can argue at length regarding the meaning of this metaphor – as, indeed, Brandt and Brandt (2005) have done.

to indicate which of two words subsequently presented (e. g., *king* or *folder*) was better related to the poetic text in which the poetic metaphor had appeared. This study demonstrates the importance of embedding the metaphor within a poetic text (even if, in this study, it was in fact a large segment of a poetic text). When the poetic metaphor appears within a text, such as in studies presenting to readers a poem for analysis (e. g., Goodblatt & Glicksohn, 2002, 2003), it is the *interanimation* of the words of the text (Richards, 1936) that enables a dynamic, bidirectional interaction (Glicksohn, 2018).

The third question is: How does one identify a metaphor within a poetic text? The major goal of various studies is to identify such metaphors (e. g., Reijnierse et al., 2018). Yet, just as there is a difference between textual foregrounding (a potentiality) and foregrounding effects (an actuality), such also is the case for the identification of poetic metaphor. As Reijnierse et al. (2018, p. 133) emphasize, "a semiotic approach can only identify cases of *potentially* deliberate metaphor.... Whether those *potentially* deliberate metaphors are psychologically real for actual language users, and in which contexts, is a question that psycholinguistic and psychological research should subsequently test, using response-elicitation approaches such as experiments or interviews." This issue has, indeed, appeared in discussions of metaphor comprehension; as Verbrugge (1980, p. 107) noted, "it is pointless to focus on a sentence and to ask whether it is or is not metaphoric. The sentence's success in occasioning a metaphoric transformation is not solely under the sentence's control." Therefore, rather than relying on the realization of a potentiality of a figurative expression as poetic metaphor, we mark such metaphors in the text for our readers (e. g., Goodblatt & Glicksohn, 2002).

Methods for Studying Poetic Metaphor and Metaphor Comprehension: Reaction Time and Rating Scales

When designing an empirical study of poetic metaphor, one must consider the experimental set-up (e. g., how the reader will view the text, how the text will be presented, whether the reader will be alone in the laboratory, whether the study will be run in a laboratory, whether the experiment will be recorded, whether physiological measurements will be recorded, etc.), as well as the task and instructions presented to the reader (e. g., whether the poetic metaphor will be indicated, what the exact instructions convey to the reader, whether the task

explicitly or implicitly suggests a comparison of texts and poetic metaphors, whether practice is required prior to the main task, whether time constraints should be employed, whether the task should be well-defined or ill-defined). The goal in this section is to present two major methods for studying poetic metaphor in the laboratory, each of which possesses a different perspective on the process of metaphor comprehension. One method relies on the *time* required by a reader to make a judgment regarding a presented metaphor, and the dependent measure here is reaction time (RT). A second method employs rating scales for making a *judgment* regarding the poetic metaphor. These two methods can be combined in the same study, enabling one to obtain a rating of a poetic metaphor on some scale (e. g., to what degree is the presented metaphor difficult or easy to comprehend) as well as the time taken to make that decision.

Turning to the first method, Hoorn (2000) has provided a useful review of the logic underlying an RT study of metaphor. Consider a poetic metaphor presented without an accompanying text, such as "Man is a machine." Would the time taken to interpret this phrase when viewed figuratively (i. e., as a metaphor) be different from the time needed to interpret a literal expression taking the same form, such as "Man is a mammal"? Would the time taken to comprehend the metaphor be dependent on the context provided, including the poetic context? This is the goal of such RT studies, and design considerations include the familiarity of the tenor and of the vehicle, the context provided (or not) for the metaphor, the duration that the metaphor is presented on the screen, the non-metaphoric phrases presented for comparison, and so forth.

Consider, for example, a study reported by Mashal et al. (2005), in which a two-word expression (e. g., "crystal river," considered by them to be "novel metaphoric, taken from poetry") was compared with a literal expression (e. g., "broken vase"). Design considerations here included a balancing of concrete and abstract words and word frequency. The participant in this study was asked "to silently decide whether the two words in the word pair are metaphorically related, literally related or unrelated" (Mashal et al., 2005, p. 2089), and the focus in this study was on comparing the patterns of brain activation while engaged in this task. In moving to a somewhat extended context, metaphoric sentences and not just two-word phrases were employed (Mashal et al., 2009).

Yet, insight regarding the meaning of a poetic metaphor takes time; hence the quality of the data generated in an empirical study of poetic metaphor is partly dependent on the conditions in which the study is run. Indeed, Gerrig (1989) has distinguished between what "can be called *time-limited comprehension*" that "is governed by the total time constraint," and "*leisurely comprehension*" that "may well involve types of processing that are largely specialized for

metaphor" (p. 238). If the reader actually verbalizes that act of metaphor comprehension, then the experimental design would have to segment the period in which a fast response is made (e. g., to make a decision that the metaphor is understood) from the subsequent verbalization of what that comprehension entails. This would also be the case when using an electrophysiological or brain-imaging technique for revealing brain activation during metaphor comprehension, because verbalization would be disruptive of any such recording. Clearly, then, empirical studies of poetic metaphor can rely on reaction time and patterns of brain activation for either acts of brief judgment or for periods of quiescence during which there is an ongoing activity of metaphor comprehension, but will not be appropriate when the reader verbalizes that process of metaphor comprehension.

Is there an intermediate option here? This is exactly where rating scales can be useful – the second major method for studying poetic metaphor, addressed in this section. Metaphors in general, and poetic metaphors in particular, can be rated on imagery (or, imageability), novelty, and appropriateness (Fainsilber & Kogan, 1984). These ratings can subsequently be used as the dependent measures of the study; they can be inter-correlated; they can be assessed for isolated poetic metaphors as well as for metaphors within poetic text. For example, is the comprehension of a poetic metaphor, such as "the sound of coming darkness" (as employed by Marks, 1982b) a function of its degree of imageability? Presumably so (Gibbs & Bogdonovich, 1999) – but maybe only for the degree of vehicle imageability. Fainsilber and Kogan (1984) found that (rated) imageability of the metaphor was positively related to the appropriateness of the metaphor, but negatively related to the novelty of the metaphor. Thus, a particular poetic metaphor might well evoke appropriate imagery, might thereby be somewhat easy to comprehend, but might also be viewed therefore as not being too novel. Degree of novelty might therefore be tied to the degree of vehicle imageability, with high novelty being tied to low imageability. Consider a study reported by Glicksohn (1994), who employed rating scales assessing the complexity of the isolated poetic metaphor and the interest aroused by the metaphor, as a function of the imageability of the tenor of the metaphor and of the imageability of the vehicle of the metaphor. High ratings in both degree of complexity and degree of interest of the metaphor were related to low ratings in degree of vehicle imageability. Complexity, interest, and novelty are ratings distinguishing between poetic metaphors. In studies such as these, such ratings can serve either as the dependent measures (or, outcomes) of the study, or as factors that need to be considered when designing an empirical study of poetic metaphor, especially of the isolated poetic metaphor.

An empirical study of poetic metaphor employing rating scales is designed to answer a theoretical question regarding poetic metaphor. Tourangeau and Rips (1991) presented three experiments; in the first experiment, tenors and vehicles were randomly paired, and the participant listed features for each. They also rated each metaphor for aptness, goodness, and comprehensibility. The features listed could be common to both tenor and vehicle, pertain to the tenor but not to the vehicle, pertain to the vehicle but not to the tenor, or those that were unique to the metaphor – *emergent* features, namely, "features not ordinarily seen as characterizing either the tenor or vehicle" (Tourangeau & Rips, 1991, p. 459). In the second experiment, participants rated each such feature, generated in the first experiment, for "its salience for the tenor and for the vehicle, its distinctiveness for the tenor and vehicle, its relationality, and its relevance to the interpretation of the metaphor" (p. 460). In the third experiment, the authors employed a set of "literary metaphors and similes from a standard anthology of modern poetry" (p. 464). For each, they "made up two interpretations – one based on features shared by the tenor and vehicle and the other based on emergent features" (p. 464), and the participants "rated the adequacy of the metaphor interpretations; they also rated the salience for the tenor and vehicle of the features comprising each interpretation" (p. 464). To exemplify this task, one of the metaphors presented was by the twentieth-century American poet William Carlos Williams: "Your knees are a southern breeze – or a gust of snow." One interpretation provided was based on common features shared by the tenor and the vehicle: "Knees are in a constant state of agitation and motion." The other interpretation provided was based on emergent features: "Knees are pleasantly surprising and thrilling." As the authors indicate, their participants "overwhelmingly preferred the interpretations based on the emergent" (p. 452).

This is but one study highlighting the existence of emergent features. Other such studies include those by Becker (1997), Gineste et al. (2000), Estes and Ward (2002), and Roncero and de Almeida (2015). The importance of emergent features has also been discussed by Wilson and Carston (2006). Of course, emergent features are exactly what one would expect to see, from a Gestalt-oriented approach to poetic metaphor. The appearance of emergent features shows that the isolated poetic metaphor is not understood based solely on the degree of preexisting similarity between its tenor and its vehicle.

Turning now to poetic metaphor appearing within a poetic text, Goodblatt (1996, 2001) employed rating scales for rating such poetic metaphors. The poems could be differentiated in terms of the number of their underlying semantic fields (Goodblatt, 1996), resulting in different degrees of interaction (which was reflected in the ratings). In the second study (Goodblatt, 2001), after completing

their ratings of the poetic metaphors, the readers prepared a written response to a series of questions regarding the poetic texts. While the written protocol provides information not readily apparent in the ratings of the metaphors (e. g., the integration of alternative readings in order to fill semantic gaps), it is the rating scale which helps to validate or operationalize a difference in poetic text and in poetic metaphor.

The synesthetic poetic metaphor was the topic of focus for Marks (1982a, 1982b). Consider "the sound of coming darkness," or "the quiet-colored end of evening" in which there is a cross-modal equivalence between loudness and brightness. Marks (1982b) asked his participants to adjust the loudness of a sound and the brightness of a light so as to best capture the cross-modal equivalence expressed in the poetic metaphor. Having established this, he could then (in a separate experiment, using different participants) employ rating scales of degree of loudness and degree of brightness. In this case, rather than using perhaps a seven-point rating scale (e. g., Goodblatt, 2001), he presented a continuous 200 mm line, that is to say, a graphic rating scale, on which, for example, the left-hand end was anchored by the term "very very soft," and the right-hand end was anchored by "very very loud." The participant's task was to mark on this scale the degree of loudness (in this case) of the poetic metaphor. Thus "the quiet-colored end of evening" would be indicated by a suitable mark on the loudness graphic rating scale, and a second suitable mark on the brightness graphic rating scale. These values then served as dependent measures of the study.

Two final issues. First, do these measures (RT and ratings) provide the researcher with valid data? Dixon and Bortolussi (2015) are wary of some studies in which "objective" measures (e. g., eye movements, reading times, brain activation indices) are collected, perhaps at the expense of the literary experience. As they write (p. 181):

> We close with a methodological comment that is intuitively obvious but often neglected in the study of literary experience. Such literary experiences may be *reactive* and change simply as a result of measurement.... For example, the effect of reading a poem in a noisy MRI scanner could well be different from the effect of reading the same poem alone in a quiet room. Moreover, the fact of measurement can lead to mental processes that would not otherwise occur.... Thus, there is always a critical need to understand the measurement and what it entails in order to further our understanding of literary experience.

Of course, one can recruit readers familiar with poetic metaphor in order to counter, somewhat, the possibly detrimental effects of the experimental set-up (see, e. g., O'Sullivan et al., 2015). Nevertheless, the question still remains: Are the "objective" measures to be preferred, or are the verbal reports to be pre-

ferred? Kuiken (2015, p. 172) makes the forceful argument that "If literary experience is the *explanandum*, validation of that pivotal construct will require verbal reports that *directly reflect* the flux of an experience self-identified as the reader's 'own.' If literary response is the *explanandum*, validation of that pivotal construct will include verbal reports that *merely complement* the neurocognitive or behavioral aspects of a complex response profile." Jacobs (2015a, p. 3) makes the equally forceful argument that "Without cross-validation by more implicit and objective measures, the results of such explicit methods may contain confounds with effects of social desirability, personal theories, or non-authentic answer elements. Moreover, not all parts of affective or cognitive responses become conscious and are verbally reportable." The implication of all of these comments cited above is that a standard RT study, in which the participant must respond as quickly as possible to the poetic metaphor presented, might provide a greatly contracted mode of processing.

The quality of the data generated is partly dependent on the conditions of the study, and partly dependent on the task presented. The measures obtained relate to both of these factors. Consider, for example, the set-up of the recent study reported by Ashby et al. (2018) concerned with reading times of nominal metaphors and similes: "Each participant sat in front of the eye tracker and used chin and forehead rests to minimize head movements.... Participants were told to read normally for comprehension, and that a question would appear after some sentences. The experiment session lasted 30–45 min" (p. 164). This is not a condition facilitating "leisurely comprehension" (Gerrig, 1989). This is also not facilitative of the type of reading encouraged by Andringa (1990, p. 232): "Although it might be possible to isolate and operationalize certain aspects, it seems at least desirable to study the potential richness and variety of reading processes in a manner as open and natural as possible." In addition, the emphasis on reading times, and not on metaphor comprehension, leaves the process uncertain. For, as Reinhart (1976, p. 395) suggested,

> A full understanding of a poetic metaphor is usually the result of the application of both focus and vehicle interpretation.... However, the relative dominance of the two procedures of understanding a given metaphor may vary. On the one hand, different speakers may tend to emphasize one of the two procedures and overlook the other. Experienced readers of poetry often emphasize the procedure of vehicle interpretation, while inexperienced readers tend to overlook this aspect of the metaphor and attempt a 'literal substitution' or interpretation of the focus.

The second issue considers the type of participant in these laboratory studies. For Marks (1982a, 1982b) none of the readers "had any special background in literature" (Marks, 1982b, p. 17). Participants in the study reported by Arzouan

et al. (2007) were students of psychology; and those reported by Mashal and Faust (2010) were "healthy participants." Presumably, in all of these studies, the reader had either very little or only minimal familiarity with poetic metaphor.

Methods for Studying Poetic Metaphor and Metaphor Comprehension: Written Protocols and Verbal Protocols

The goal in this section is to present two more major methods for studying poetic metaphor: the written protocol and the verbal protocol. These can be used either in the laboratory or in the field. Both of these will result in a text, provided by the reader, describing metaphor comprehension. Addressing the first method, written protocols can be obtained in various ways: readers can provide a written response to a series of questions regarding the poetic metaphor and poetic text (e. g., Goodblatt, 2001); they can generate a short, written interpretation of poetic metaphors (Gibbs & Bogdonovich, 1999; Shen 2008); they can select lines from a text they find to be striking or evocative, and then make verbal comments on these (e. g., Sikora et al., 2011). In fact, going back to IA Richards's pioneering empirical study of poetic metaphor and poetic texts, readers can "spend a certain amount of time...reading them [a sheet of poems] and writing short comments upon them" (Richards, 1925, pp. 1–2; see Goodblatt & Glicksohn, 2010, for further discussion of this notebook). These "protocols," as Richards termed them, were limited to "100 words to each poem or passage" (Richards, 1925, pp. 1–2), and they were to be the result of the readers having spent "a certain amount of time – it needn't be very long and in any case it will be – I may perhaps dare to say – remarkably well spent time" in reading and commenting on the texts.

The second method refers to the transcription of a verbal report, provided by a reader of a poetic text. Consider a poetic metaphor presented without accompanying text, such as "Man is a machine." What does this mean? Instead of thinking about this in silence, one can think out aloud, in real time. That verbal report of the process of metaphor comprehension is recorded, and subsequently transcribed. The transcribed text, which is the verbal protocol, is then amenable to detailed content analysis. Peskin (1998) reported a study wherein "experts" (PhD candidates in English) and "novices" (high-school or undergraduate students) thought aloud "as they attempted to make sense of 2 period poems" (Pe-

skin, 1998, p. 235). Two factors clearly distinguishing between the verbal protocols of these two groups were: (a) the degree of knowledge brought to the task; and (b) the type of strategy employed. Instructions presented to readers for thinking aloud will also influence the verbal protocol. In Eva-Wood's (2004) study, one group of readers was given instructions to "think aloud" while reading a poem, and a second group was asked to "think-and-feel aloud." In the Sikora et al. (2011) study, readers were asked to describe their experience "in as much detail as possible" (p. 260) of an evocative section of a poem that they had previously selected. This, then, is a variant of the "think-and-feel aloud" instruction.

Goodblatt and Glicksohn (2002, 2003, 2010) have employed thinking aloud in their studies, coupled with a depiction of the hypothesized reading process presented as a flowchart. In these studies, poetic metaphors within text appear in bold for the readers to engage with. The process of metaphor comprehension is viewed in terms of problem solving, each poetic metaphor presenting for the reader a problem to solve. For example, in the Goodblatt and Glicksohn (2002) study, readers were presented with a poem ("Arrival," by William Carlos Williams) containing the line "dropping **its silk and linen leaves**." Two alternative readings were suggested in the flowchart for this metaphoric phrase: (a) a dress (that term appearing previously in the poem) being perceived as autumn leaves; and/or (b) the autumn (that term also appearing previously in the poem) leaves being perceived as a dress. In one verbal protocol cited, a reader focused on the semantic anomaly: "leaves cannot be silk and linen…therefore it's metaphor, but they're beautiful so the poet compares them to silk and linen." For this reader, the problem presented by the poetic metaphor is "resolved" just by acknowledging the semantic anomaly. A second protocol cited exhibits extended problem solving: "the fact that her dress is described as silk and linen leaves that drop about her ankles … it could be, maybe she's Mother Nature? She's a woman, nature is a woman, and she is adorned in silk and linen leaves. Beauty of nature." Janssen et al., (2003) elaborated on this process of probem solving (and problem *finding*, which in itself is of interest for empirical studies of poetic metaphor). They suggest three categories that would be indicative of problem solving in the verbal protocol. Examples from the above protocols exemplify the following types of response: (a) detecting a problem (e. g., "leaves cannot be silk and linen"); (b) solving that problem (e. g., "maybe she's Mother Nature?"); and (c) responding meta-cognitively (which, as they suggest, would be an indication in the protocol that the reader is commenting on the reading process, such as, "let me just read that phrase again").

Following Hauff and Fogarty (1996), "good" readers should be asking more questions about the text (i. e., detecting problems and then trying to solve

them), while "poor" readers should spend more time simply reading and re-reading the poetic text. That can be clearly seen in the protocols obtained in our studies. In fact, the Goodblatt and Glicksohn (2010) paper reported a comparison in terms of the length of the verbal protocol: readers exhibiting pronounced problem solving in their verbal protocols provided longer texts than those not engaged in problem solving. The verbal protocol is clearly a rich source of data for the empirical study of poetic metaphor.

Two issues are addressed in the present discussion of studies employing RT and rating scales. The first issue is whether protocols (written or verbal) can provide the researcher with valid data. Martindale (Martindale, 2008; Martindale & Dailey, 1995) has proposed that Richards's research is problematic, arguing, in effect, that Richards himself completely misunderstood his data. This is because he neither used rating scales nor statistics for analyzing the degree of agreement among the various readers regarding the poems presented to them. The use of rating scales and their statistical analysis is clearly one viable option for studying poetic texts, as has been previously discussed in this chapter (e. g., Goodblatt, 1996). This is, however, not the only method available for studying such texts (e. g., Goodblatt, 2001). What is more, research that generates protocols, such as used by Richards, must surely illuminate aspects of the reading process not readily assessed using other methods such as rating scales. This is indeed noted by Martindale and Dailey (1995, p. 307), when they write:

> The most likely response of a literary critic to our results would be that we have demonstrated agreement that is not on a level that literary critics care about.... It is true that we have not demonstrated that people make the same symbolic or deep-level interpretations of poetry. However, such interpretations require a theoretical background that the ordinary reader does not have.

Martindale and Dailey rightly note that literary scholars are interested in questions that their type of research does not investigate – and that is a focus of the type of research pioneered by Richards. In addition, the concern of this research is the exceptional reader, just as the investigation of problem solving among chess masters is reserved for this select group. The preferred reader for understanding a poetic text is thus one who has an adequate background for engaging with the text and its context – and this is very likely a student of literature.

Certain precautions should be taken when using a thinking-aloud method. In Warren's (2011) study, the readers were given instructions indicating that "Many people mumble comments to themselves when they read or write – the purpose of a think-aloud session is simply to raise the volume of your mumbling. Don't censor anything.... Don't explain or justify what you are doing" (p. 354). Andringa (1990, p. 235) reports that "the procedure of thinking aloud

after whole paragraphs does not disturb the reading process as much as thinking aloud sentence after sentence. However, some direct response might get lost or be supplanted." Fox et al. (2011, p. 337) acknowledge that "the time-consuming collection and analysis of think-aloud protocols will provide important insights to warrant the investment of time and resources" – but with a focus on tasks having "large stable individual differences in performance." The empirical study of poetic metaphor is just such a task.

Following Pressley and Afflerbach (1995, p. 126), but with reference to poetic metaphor, a productive approach here would be: (a) to provide a flowchart of the "processing that might be expected in the particular situation"; (b) provide a detailed analysis of the verbal protocol; and (c) the "collection of behavioral measures, such as objective memory of text, reading times, and so on." These are all viable objectives of an empirical study of poetic metaphor. One could, further, add the collection of eye movements, as the reader engages with the text (de Vries et al., 2018; Lauwereyns & d'Ydewalle, 1996), but as Lauwereyns and d'Ydewalle (1996) readily note, "Eye movements as the unique source of information for knowledge acquisition would mostly be uninterpretable, unless eye movements are explicitly used as responses with a predefined meaning" (p. 16). In their study, these authors combined eye movements with a thinking-aloud protocol, as an "expert" analyzed a poem, thinking aloud while eye movements were recorded (the verbal protocol being recorded as the soundtrack of the video recording of the eye movements). Note that these same considerations are pertinent for studies of poetic metaphor employing electrophysiological recording or functional brain imaging: brain activity must be tied to cognitive activity in order for the interpretation of one source of data to be supported by the interpretation of the other source of data. And, as with all empirical studies (including those of poetic metaphor), the actual process of collecting the data might be disruptive of the process under study (here, the process of metaphor comprehension).

The second issue considers the type of participant in these studies generating such protocols. Thus for Nenadić, et al. (2019), "expert" readers were "comprised of poets (with at least two published books), literary critics, or university professors of literature or languages whose work is related to the field of poetry." Biermann (1997, p. 63) suggests that "it would be a valuable contribution to the field of literary studies if empirical research could undertake the task of refining the highly sophisticated existing literary explanations based on principles that have been derived from the insights of such talented professional readers." For Goodblatt (1996, 2001), the readers were advanced students of English language and literature. In these studies, the reader possessed a certain

degree of expertise in reading poetic metaphor, from prior or concurrent study, writing or teaching.

Waggoner (1990, p. 105) argues that "Future research on metaphor, in particular, research aimed at investigating the claims of interaction theory, should address the comprehension of elite metaphors by individuals who deal with them on a regular basis." In line with Waggoner, poetic metaphor requires the mind of a competent reader (in whatever sense) who will be able to analyze the interaction of its parts (what we refer to as an act of problem solving – see Goodblatt & Glicksohn, 2002), while being appreciative of the aesthetic experience of reading the poetic metaphor within its poetic context. Note that this aesthetic experience can include one of acknowledging the *grotesque* nature of a poetic metaphor and its text (see Goodblatt & Glicksohn, 2017b), or of the *sublime* nature of the poetic metaphor and its text (see Miall, 2007).

In Warren's (2011) study, the reader is a professor of English. Yet as Warren notes, "English professors, just like experts in other disciplines, develop both "generic" and "specific" expertise" (p. 349). To investigate these types of expertise in reading poetic metaphor, the readers read poems that were either well within their specialty, or far from their specialty, and provided a verbal protocol for each. It is certainly enlightening to look at two such protocols in reading the poem "The Flea," written by the seventeenth-century English Metaphysical poet John Donne. At the center of this poem is the Metaphysical conceit, defined concisely by Helen Gardner: "A conceit is a comparison whose ingenuity is more striking than its justness, or, at least, is more immediately striking. All comparisons discover likeness in things unlike: a comparison becomes a conceit when we are made to concede likeness while being strongly conscious of unlikeness" (Gardner, 1972/1957, p. 19). In this instance, the conceit as metaphor – "This flea is you and I" – extends the comparison between, on the one hand, the mingling of the two lovers' bloods within the body of the flea that has bitten them, and, on the other hand, the physical and spiritual union of the lovers through the sexual act. Juxtaposing passages from the protocols of two very different readers demonstrates the way in which useful protocols are most effectively produced for the study of poetic metaphor.

The first reader, who is a scholar of Renaissance literature, states:

> First of all, a very familiar poem [Donne's "The Flea"], so all the obvious things would be things I guess one would not want to say in the session [note: the task presented to the reader was to think aloud and to compose a short text proposing a hypothetical talk for a conference]. That's the first thing I'm thinking of: what can I say that's new? Everybody knows the conceit; everybody knows how it works; everybody knows the bloods mingle inside the flea; everybody knows that it's kind of a *carpe diem* [seize the day] seduction poem. (Warren, 2011, p. 358)

The second reader, a scholar of late nineteenth- and early twentieth-century literature, states:

> I mean I guess I could bring "The Flea" into this discussion to mediate between the spiritualized sexuality of Hopkins and the cheapened sexuality I find in Eliot. I think Eliot on the erotic is awful – there's no such thing as good sex in Eliot's corpus. Whereas "The Flea" I think is a pragmatic, sensible, playful invocation of the erotic between two poles. I just worry about being a historically irresponsible critic. (Warren, 2011, p. 360)

Short and selective as these examples are, the difference in reading this metaphor is clearly apparent. One criterion, then, is established: the need for a competent reader, in other words one who possesses previous experience in the analysis of metaphor – in this instance, specifically in relation to the conceit. Thus the first reader distinctly marks the conceit, discussing the physical basis in its comparison of a flea bite with the lovers' sexual act. In contrast, the second reader does not acknowledge the existence of the conceit in the poem. A second criterion is also established: the need for an understanding of a specific tradition of poetry (and consequently of poetic metaphor). Thus while it is evident that the first reader is familiar with the tradition of Metaphysical conceit, the second reader turns instead to a more general discussion of sexuality, positioning Donne's poem with respect to poetic traditions of his area of scholarship (the English Victorian poet George Manley Hopkins, 1844–1889; the Anglo-American Modernist poet T. S. Eliot, 1888–1965).

What is to be Gained from an Empirical Study of Poetic Metaphor?

The *empirical* study of poetic metaphor can make three contributions to the literature beyond that of a *literary* analysis of poetic metaphor. The first is to create a bridge between literary scholarship and cognitive psychology. The second is the possibility for studying the *process* of metaphor comprehension, while retaining experimental control. Finally, an empirical study of poetic metaphor enables a hypothesis of theoretical interest to be submitted to test.

The necessity for a bridge between literary scholarship and cognitive psychology was noted twenty-five years ago by Gibbs (1994, pp. 260–261) who writes:

> One of the disappointing aspects of the psychological research on metaphor understanding is that few studies have been devoted to literary metaphor.... One of the challenges for

> psychologists and others interested in the empirical study of metaphor is to find ways of examining how it is that readers make sense of novel metaphorical texts. Doing so requires that we go beyond the isolated metaphor to envision better how metaphors are recognized as intentionally created by authors to make new the world we live in.

More recently, Holyoak and Stamenković (2018, pp. 658–659) came to a similar conclusion. They suggested that:

> By broadening the range of metaphors considered by psychologists, it may be possible to contribute to advances in the closely related field of literary psychology.... A number of unresolved issues might be usefully explored further in connection with literary metaphors.... literary metaphors often seem ambiguous and open-ended.... The metaphors and symbols found in poetry may have implications for the creation and comprehension of novel metaphors and also the acquisition of relatively universal symbols.

Such a bridge between literary scholarship and cognitive psychology can be found in the work of those research collaborations of a literary scholar and a cognitive psychologist. Thus Miall and Kuiken (1994, 1998) update Shklovsky's notion of "art as device" to investigate readers' affective responses to literary texts. On their part, Goodblatt and Glicksohn (2003, 2010) update I. A. Richards's empirical technique of obtaining written protocols to one of obtaining online verbal protocols of metaphor comprehension. This chapter will therefore make an impact if, indeed, cognitive psychologists pay attention to the way *poetic metaphor* has been studied, while literary scholars consider how *real readers* read the texts.

A second contribution to be gained from an empirical study of poetic metaphor resides in the possibility for studying the process of metaphor comprehension, while retaining experimental control. Pressley and Afflerbach (1995, pp. 128–129) note that "the manner in which text is presented to the reader can influence the nature of the reporting." The Goodblatt and Glicksohn (2002, 2003, 2010) studies employed a microgenetic procedure, wherein more and more text appears on the screen. A microgenetic technique can be used to display images (Glicksohn & Yaniv, 2016) or to present poetic texts (Goodblatt & Glicksohn, 2002), and in each case the verbal report can be analyzed with this specific order of presentation of image or text taken into consideration. In the Mashal and Faust (2010) study, four-line poems were displayed on screen, followed by a probe word, described to the participants as "words that best described the theme of the text" (p. 24). The participant in that study was required "to read and understand...[the text] silently and to decide as quickly as possible whether the following probe word had appeared or had not appeared in the text by pressing one of two response buttons" (p. 24). The authors suggest that "de-

ciding whether the probe word best described the theme of the text required that the participants read and understand the text at a level that encouraged the integration of the lines into a coherent representation" (p. 25). Note, therefore, that in all of these examples, there is experimental control within the task.

Finally, an empirical study of poetic metaphor will enable one to test a hypothesis of theoretical interest. Consider, for example, the relationship between Interaction Theory and Blending Theory. Following the groundwork set in place by Black (1962) in proposing the Interaction Theory of metaphor, the process of metaphor comprehension has at least three stages in its development (Goodblatt & Glicksohn, 2016, 2017a). The tenor (or target, or primary subject) is first seen through the lens of the vehicle (or source, or secondary subject); in the second stage, the vehicle is now seen through the lens of an already transformed tenor; in the third stage, a transformed tenor is now seen through the lens of a transformed vehicle, as well as a transformed vehicle being seen through the lens of a transformed tenor. Thus, both tenor and vehicle "are shifting percepts or concepts within a dynamic, interactive process. Bidirectionality is an integral part of this interactive process" (Goodblatt & Glicksohn, 2017a, p. 3), and this notion has a venerable history. Franklin (1988, p. 161), for example, in referring *inter alia* to Black's (1962) Interaction Theory of metaphor, notes that "When abstract works have titles that refer to less tangible states of afairs ... the vectors of structuring are bidirectional: as the visual image is configured to show '*agony*', particular connotations of the word come to the fore. In such cases, the achievement of fit between title and artwork meanings is often mediated by expressive or 'physiognomic' properties." Other writers expressing views in support of bidirectionality include Jackendoff and Aaron (1991, pp. 334–335), Goldwasser (1999, p. 610), and Indurkhya (2006, p. 158).

Given the convergence between our Gestalt-oriented appoach to poetic metaphor and that of blending theory (Fauconnier & Turner, 1998; Vandaele & Brône, 2009, p. 12), in what respect do the two approaches differ? Fauconnier and Turner (1998, pp. 138–140) argue that "The blend contains emergent structure not in the inputs...Second, *completion* brings additional structure to the blend...By *completion*, this familiar structure is recruited into the blend." Thus the blend is the emergent whole. As Gibbs (2009, p. 28) has argued, however, "blending theory has had no significant impact on psycholinguistic studies of verbal metaphor understanding ... One problem is that blending theory offers descriptions of idealized speaker/listeners' full-blown understandings of language yet does not provide specific, detailed, unique hypotheses about understanding processes that can be tested readily in the laboratory." The notion of bidirectionality includes the possibility that there is no such blending. Rather both unidirectional readings are preserved in the same verbal protocol, a hy-

pothesis that has been submitted to empirical investigation (e. g., Goodblatt & Glicksohn, 2017b).

Bruhn (2018) has recently suggested that it is the source (vehicle) which is "created and construed" (p. 713) according to the target (tenor). Glicksohn (2018), in turn, suggests that Bruhn (2018) might well be correct – but not with respect to metaphor *comprehension*, rather to metaphor *production*. This now becomes an issue for empirical investigation. If Bruhn (2018) is correct, then one should be able to find evidence for the target coming first, and the source being constructed appropriately in a verbal protocol of metaphor comprehension. If Glicksohn (2018) is correct, then one should be able to find evidence for the source coming first, and the target being viewed appropriately in that verbal protocol. One could also design an RT study to test these hypotheses, although that type of study might only be able to reveal which term of the poetic metaphor is dominant. What a "source being construed appropriately" (Bruhn, 2018) or a "target being viewed appropriately" (Glicksohn, 2018) might entail more than an RT study could possibly provide. One could also explore these hypotheses for the case of metaphor production, but that does not come under the scope of the present chapter.

References

Andringa, E. (1990). Verbal data on literary understanding: A proposal for protocol analysis on two levels. *Poetics, 19*(3), 231–257. https://doi.org/10.1016/0304-422X(90)90022-W

Aryani, A., Kraxenberger, M., Ullrich, S., Jacobs, A. M., & Conrad, M. (2016). Measuring the basic affective tone of poems via phonological saliency and iconicity. *Psychology of Aesthetics, Creativity, and the Arts, 10*(2), 191–204. https://doi.org/10.1037/aca0000033

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Dynamics of hemispheric activity during metaphor comprehension: Electrophysiological measures. *NeuroImage, 36*(1), 222–231. https://doi.org/10.1016/j.neuroimage.2007.02.015

Ashby, J., Roncero, C., de Almeida, R. G., & Agauas, S. J. (2018). The early processing of metaphors and similes: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology, 7*(1), 151–168. https://doi.org/10.1080/17470218.2016.1278456

Bambini, V., Canal, P., Resta, D., & Grimaldi, M. (2019). Time course and neurophysiological underpinnings of metaphor in literary context. *Discourse Processes, 56*(1), 77–97. https://doi.org/10.1080/0163853X.2017.1401876

Becker, A. H. (1997). Emergent and common features influence metaphor interpretation. *Metaphor and Symbol*, *12*(4), 243–259. https://doi.org/10.1207/s15327868ms1204_3

Biermann, I. (1997). When metaphor counts. *Language and Literature, 6*, 57–68. https://doi.org/10.1177/096394709700600104

Black, M. J. (1962). *Models and metaphors: Studies in language and philosophy*. Cornell UP.

Brandt, L., & Brandt, P. A. (2005). Making sense of a blend: A cognitive-semiotic approach to metaphor. *Annual Review of Cognitive Linguistics, 3*(1), 216–249. https://doi.org/10.1075/arcl.3.12bra

Bruhn, M. J. (2018). Target first: On "bidirectionality and metaphor." *Poetics Today*, *39*(4), 703–733. https://doi.org/10.1215/03335372-7150952

de Vries, C., Reijnierse, W. G., & Willems, R. M. (2018). Eye movements reveal readers' sensitivity to deliberate metaphors during narrative reading. *Scientific Study of Literature*, *8*(1), 135–164. https://doi.org/10.1075/ssol.18008.vri

Dixon, P., & Bortolussi, M. (2015). Measuring literary experience: Comment on Jacobs (2016). *Scientific Study of Literature, 5*(2), 178–182. https://doi.org/10.1075/ssol.5.2.03dix

Estes, Z., & Ward, T. B. (2002). The emergence of novel attributes in concept modification. *Creativity Research Journal, 14*(2), 149–156. https://doi.org/10.1207/S15326934CRJ1402_2

Eva-Wood, A. L. (2004). Thinking and feeling poetry: Exploring meanings aloud. *Journal of Educational Psychology, 96*(1), 182–191. https://doi.org/10.1037/0022-0663.96.1.182

Fainsilber, L., & Kogan, N. (1984). Does imagery contribute to metaphoric quality? *Journal of Psycholinguistic Research, 13*(5), 383–391. https://doi.org/10.1007/BF01068153

Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science, 22*(2), 133–187. https://doi.org/10.1016/S0364-0213(99)80038-X

Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M. J., Pléh, C., & Jacobs, A. M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words. *NeuroImage, 63*(3), 1432–1442. https://doi.org/10.1016/j.neuroimage.2012.07.029

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137*(2), 316–344. https://doi.org/10.1037/a0021663

Franklin, M. B. (1988). "Museum of the Mind": An inquiry into the titling of artworks. *Metaphor and Symbolic Activity, 3*(3), 157–174. https://doi.org/10.1207/s15327868ms0303_4

Franklin, M. B. (1994). A feeling for words: Arnheim on language. *The Arts in Psychotherapy, 21*(4), 261–267. https://doi.org/10.1016/0197-4556(94)90005-1

Gambino, R., & Pulvirenti, G. (2019). Neurohermeneutics: A transdisciplinary approach to literature. *Gestalt Theory*, *41*(2), 185–200. https://doi.org/10.2478/gth-2019-0018

Gardner, H. (1972/1957). Introduction. In H. Gardner (Ed.), *The metaphysical poets* (pp. 15–29). Penguin.

Gentner, D., Bowdle, B. F., 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). MIT Press.

Gerrig, R. J. (1989). Empirical constraints on computational theories of metaphor: Comments on Indurkhya. *Cognitive Science, 13*(2), 235–241. https://doi.org/10.1016/0364-0213(89)90005-0

Gibbs, R. W., Jr. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge UP.

Gibbs, R. W. (2009). Why do some people dislike conceptual metaphor theory? *Cognitive Semiotics, 5*(1-2), 14–36. https://doi.org/10.1515/cogsem.2013.5.12.14

Gibbs, R. W., Jr., & Bogdonovich, J. (1999). Mental imagery in interpreting poetic metaphor. *Metaphor and Symbol, 14*(1), 37–54. https://doi.org/10.1207/s15327868ms1401_4

Gineste, M.-D., Indurkhya, B., & Scart, V. (2000). Emergence of features in metaphor comprehension. *Metaphor and Symbol, 15*, 117–135. https://doi.org/10.1207/S15327868MS1503_1

Glicksohn, J. (1994). Putting interaction theory to the empirical test: Some promising results. *Pragmatics & Cognition, 2*, 223–235. https://doi.org/10.1075/pc.2.2.02gli

Glicksohn, J. (2018). Bidirectionality and interaction in metaphor comprehension and metaphor production: A reply to Mark J. Bruhn. *Poetics Today*, *39*(4), 735–739. https://doi.org/10.1215/03335372-7150966

Glicksohn, J., & Berkovich-Ohana, A. (2012). Absorption, immersion, and consciousness. In J. Gackenbach (Ed.), *Video game play and consciousness* (pp. 83–99). New York: Nova Science Publishers, Inc.

Glicksohn, J., & Goodblatt, C. (1993). Metaphor and Gestalt: Interaction theory revisited. *Poetics Today, 14*(1), 83–97. https://doi.org/10.2307/1773141

Glicksohn, J., & Yaniv, H. (2016). Visual hybrids induce anxiety: A microgenetic approach. *Psychology of Consciousness: Theory, Research, and Practice, 3*(3), 239–257. https://doi.org/10.1037/cns0000085

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97*(1), 3–18. https://doi.org/10.1037/0033-295X.97.1.3

Goldwasser, O. (1999). Book review: *Pictorial metaphor in advertising*. *Journal of Pragmatics, 31*, 609–618. https://doi.org/10.1016/S0378-2166(98)00080-0

Goodblatt, C. (1996). Semantic fields and metaphor: Going beyond theory. *Empirical Studies of the Arts, 14*, 65–78. https://doi.org/10.2190/GQP6-TP8K-1J1Y-0JP7

Goodblatt, C. (2001). Adding an empirical dimension to the study of poetic metaphor. *Journal of Literary Semantics, 30*(3), 167–180. https://doi.org/10.1515/jlse.2001.008

Goodblatt, C., & Glicksohn, J. (2002). Metaphor comprehension as problem solving: An online study of the reading process. *Style, 36*(3), 428–445. https://www.jstor.org/stable/10.5325/style.36.3.428

Goodblatt, C., & Glicksohn, J. (2003). From *Practical Criticism* to the practice of literary criticism. *Poetics Today, 24*(2), 207–236. https://doi.org/10.1215/03335372-24-2-207

Goodblatt, C., & Glicksohn, J. (2010). Conversations with I. A. Richards: The renaissance in cognitive literary studies. *Poetics Today, 31*(3), 387–432. https://doi.org/10.1215/03335372-2010-001

Goodblatt, C., & Glicksohn, J. (2016). Bidirectionality in poetic metaphor: William Carlos Williams and Imagist poetry. *Versus. Quaderni di studi semiotici, 122*(1), 93–110. https://doi.org/10.14649/84880

Goodblatt, C., & Glicksohn, (2017a). Bidirectionality and metaphor: An introduction. *Poetics Today*, *38*(1), 1–14. https://doi.org/10.1215/03335372-3716189

Goodblatt, C., & Glicksohn, J. (2017b). *Discordia concors* and bidirectionality: Embodied cognition in John Donne's *Songs and Sonnets*. *Poetics Today, 38*(1), 163–188. https://doi.org/10.1215/03335372-3716304

Gräbe, I. (1984). Local and global aspects of interaction processes in poetic metaphor. *Poetics, 13*(6), 433–457. https://doi.org/10.1016/0304-422X(84)90017-2

Hauff, H. M., & Fogarty, G. J. (1996). Analysing problem solving behaviour of successful and unsuccessful statistics students. *Instructional Science, 24*(6), 397–409. https://doi.org/10.1007/BF00125577

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin, 144*, 641–671. https://doi.org/10.1037/bul0000145

Hoorn, J. F. (2000). The hazard of hidden interactions: A reanalysis of designs in reaction-time studies on metaphor. *CLCWeb: Comparative Literature and Culture, 2*(3), article 4. https://doi.org/10.7771/1481-4374.1077

Indurkhya, B. (2006). Emergent representations, interaction theory and the cognitive force of metaphor. *New Ideas in Psychology, 24*, 133–162. https://doi.org/10.1016/j.newideapsych.2006.07.004

Jackendoff, R., & Aaron, D. (1991). *More Than Cool Reason: A Field Guide to Poetic Metaphor* by George Lakoff; Mark Turner. *Language, 67*(2), 320–338. https://doi.org/10.2307/415109

Jacobs, A. M. (2015a). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience, 9*, *186*. https://doi.org/10.3389/fnhum.2015.00186

Jacobs, A. M. (2015b). The scientific study of literary experience: Sampling the state of the art. *Scientific Study of Literature, 5*(2), 139–170. https://doi.org/10.1075/ssol.5.2.01jac

Jacobs, A. M., & Kinder, A. (2018). What makes a metaphor literary? Answers from two computational studies. *Metaphor and Symbol, 33*(2), 85–100. https://doi.org/10.1080/10926488.2018.1434943

Janssen, T., Braaksma, M., Rijlaarsdam, G., & Couzijm, M. (2003, August 26–30). *Reading for points and problems: A study of students' literary interpretation processes.* Paper presented at the EARLI conference, Padova, Italy.

Katz, A. N., Paivio, A., & Marschark, M. (1985). Poetic comparisons: Psychological dimensions of metaphoric processing. *Journal of Psycholinguistic Research, 14*(4), 365–383. https://doi.org/10.1007/BF01067881

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*(4), 191–214. https://doi.org/10.1207/s15327868ms0304_1

Kenett, Y. N., Gold, R., & Faust, M. (2018). Metaphor comprehension in low and high creative individuals. *Frontiers in Psychology, 9*, 482. https://doi.org/10.3389/fpsyg.2018.00482

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review, 7*(2), 257–266. https://doi.org/10.3758/BF03212981

Kittay, E. F. (1997). Of "men" and metaphors: Shakespeare, embodiment and filing cabinets. In T. N. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 375–402). American Psychological Association.

Kuiken, D. (2015). The implicit erasure of "literary experience" in empirical studies of literature: Comment on "The Scientific study of literary experience: Sampling the state of the art" by Arthur Jacobs. *Scientific Study of Literature, 5*(2), 171–177. https://doi.org/10.1075/ssol.5.2.02kui

Lauwereyns, J., & d'Ydewalle, G. (1996). Knowledge acquisition in poetry criticism: The expert's eye movements as an information tool. *International Journal of Human-Computer Studies, 45*(1), 1–18. https://doi.org/10.1006/ijhc.1996.0039

Marks, L. E. (1982a). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8*(2), 177–193. https://doi.org/10.1037/0096-1523.8.2.177

Marks, L. E. (1982b). Synesthetic perception and poetic metaphor. *Journal of Experimental Psychology: Human Perception and Performance, 8*(1), 15–23. https://doi.org/10.1037/0096-1523.8.1.15

Martindale, C. (2008). Response: psychological foundations of literary theory. *Journal of Literary Theory*, *1*(2), 447–457. https://doi.org/10.1515/JLT.2007.027

Martindale, C., & Dailey, A. (1995). I. A. Richards revisited: Do people agree in their interpretations of literature? *Poetics, 23*(4), 299–314. https://doi.org/10.1016/0304-422X(94)00025-2

Mashal, N., & Faust, M. (2010). The effects of metaphoricity and presentation style on brain activation during text comprehension. *Metaphor and Symbol, 25*(1), 19–33. https://doi.org/10.1080/10926480903538464

Mashal, N., Faust, M., & Hendler, T. (2005). The role of the right hemisphere in processing nonsalient metaphorical meanings: Application of principal components analysis to fMRI data. *Neuropsychologia, 43*(14), 2084–2100. https://doi.org/10.1016/j.neuropsychologia.2005.03.019

Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2009). An fMRI study of processing novel metaphoric sentences. *Laterality, 14*(1), 30–54. https://doi.org/10.1080/13576500802049433

McCabe, A. (1983). Conceptual similarity and the quality of metaphor in isolated sentences versus extended contexts. *Journal of Psycholinguistic Research, 12*(1), 41–68. https://doi.org/10.1007/BF01072713

Miall, D. S. (2007). Foregrounding and the sublime: Shelley in Chamonix. *Language and Literature*, *16*(2), 155–168. https://doi.org/10.1177/0963947007075982

Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics, 22*(5), 389–407. https://doi.org/10.1016/0304-422X(94)00011-5

Miall, D. S., & Kuiken, D. (1998). The form of reading: Empirical studies of literariness. *Poetics, 25*(6), 327–341. https://doi.org/10.1016/S0304-422X(98)90003-1

Nenadić, F., Vejnović, D., & Marković, S. (2019). Subjective experience of poetry: Latent structure and differences between experts and non-experts. *Poetics*, *73*, 100–113. https://doi.org/10.1016/j.poetic.2018.11.005

O'Sullivan, N., Davis, P., Billington, J., Gonzalez-Diaz, V., & Corcoran, R. (2015). "Shall I compare thee": The neural basis of literary awareness, and its benefits to cognition. *Cortex, 73*, 144–157. https://doi.org/10.1016/j.cortex.2015.08.014

Peskin, J. (1998). Constructing meaning when reading poetry: An expert-novice study. *Cognition and Instruction, 16*(3), 235–263. https://doi.org/10.1207/s1532690xci1603_1

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Lawrence Erlbaum

Reijnierse, W. G., Burgers, C., Krennmayr, T., & Steen, G. J. (2018). DMIP: A method for identifying potentially deliberate metaphor in language use. *Corpus Pragmatics, 2*(2), 129–147. https://doi.org/10.1007/s41701-017-0026-7

Reinhart, T. (1976). On understanding poetic metaphor. *Poetics, 5*(4), 383–402. https://doi.org/10.1016/0304-422X(76)90017-6

Richards, I. A. (1925). *Practical Criticism – Notebook 1925*.

Richards, I. A. (1936). *The philosophy of rhetoric*. Oxford UP.

Roncero, C., & de Almeida, R. G. (2015). Semantic properties, aptness, familiarity, conventionality, and interpretive diversity scores for 84 metaphors and similes. *Behavior Research Methods, 47*(3), 800–812. https://doi.org/10.3758/s13428-014-0502-y

Shen, Y. (2008). Metaphor and poetic figures. In R. W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 295–307). Cambridge UP.

Sikora, S., Kuiken, D., & Miall, D. S. (2011). Expressive reading: A phenomenological study of readers' experience of Coleridge's *The Rime of the Ancient Mariner*. *Psychology of Aesthetics, Creativity, and the Arts, 5*(3), 258–268. https://doi.org/10.1037/a0021999

Snævarr, S. (2010). *Metaphors, narratives, emotions: Their interplay and impact*. Rodopi B. V.

Stamenković, D., Ichien, N., & Holyoak, K. J. (2019). Metaphor comprehension: An individual-differences approach. *Journal of Memory and Language, 105*, 108–118. https://doi.org/10.1016/j.jml.2018.12.003

Tourangeau, R., & Rips, L. (1991). Interpreting and evaluating metaphors. *Journal of Memory and Language, 30*(4), 452–472. https://doi.org/10.1016/0749-596X(91)90016-D

Vandaele, J., & Brône, G. (2009). Cognitive poetics: A critical introduction. In G. Brône & J. Vandaele (Eds.), *Cognitive poetics: Goals, gains and gaps* (pp. 1–29). Mouton de Gruyter.

Van Peer, W., Hakemulder, J., & Zyngier, S. (2007). Lines on feeling: Foregrounding, aesthetics and meaning. *Language and Literature, 16*(2), 197–213. https://doi.org/10.1177/0963947007075985

Verbrugge, R. R. (1980). Transformations in knowing: A realist view of metaphor. In R. P. Honeck & R. R. Hoffman (Eds.), *Cognition and figurative language* (pp. 87–125). Lawrence Erlbaum

Waggoner, J. E. (1990). Interaction theories of metaphor: Psychological perspectives. *Metaphor and Symbolic Activity, 5*(2), 91–108. https://doi.org/10.1207/s15327868ms0502_3

Warren, J. E. (2011). "Generic" and "specific" expertise in English: An expert/expert study in poetry interpretation and academic argument. *Cognition and Instruction, 29*(4), 349–374. https://doi.org/10.1080/07370008.2011.607929

Wertheimer, M. (2010). A Gestalt perspective on the psychology of thinking. In B. Glatzeder, V. Goel, & A. Müller (Eds.), *Towards a theory of thinking* (pp. 49–58). Springer.

Wilson, D., & Carston, R. (2006). Metaphor, relevance and the 'emergent property' issue. *Mind & Language, 21*(3), 404–433. https://doi.org/10.1111/j.1468-0017.2006.00284.x

Zyngier, S., Van Peer, W., & Hakemulder, J. (2007). Complexity and foregrounding: In the eye of the beholder? *Poetics Today, 28*(4), 653–682. https://doi.org/10.1215/03335372-2007-011