What's it Take to Get a Novel Metaphor Around Here? Exploring Potential Contributors to Novel Metaphor Production

A Thesis Submitted to the Committee on Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science in the Faculty of Arts and Science

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Psychology M.Sc. Graduate Program

September 2025

Abstract

What's it Take to Get a Novel Metaphor Around Here? Exploring Potential Contributors to Novel Metaphor Production

Kiera Johnston

Metaphors are a ubiquitous component of communication, which are either conventional (familiar) or novel phrases. Metaphors are useful for discussing abstract concepts such as emotions that can be difficult to discuss literally, and as such benefit social relationships (Nippold et al., 2017). Research on contributors to novel metaphor production is limited and previous methodologies have been flawed, however some research suggests a role for executive function (EF; e.g., Menashe et al., 2020). Additionally, figurative language comprehension and metaphor production have been demonstrated to rely on overlapping abilities (Benedek et al., 2014). The current study examined figurative language comprehension and EF in novel metaphor production in 112 young adults aged 17-29 years. Neither individual nor groupwise analyses revealed any relations between EF or figurative language comprehension and novel metaphor production. However, an association with nonverbal IQ was found, suggesting that abstract reasoning abilities may play an important role in metaphoric ability.

Keywords: Novel metaphor, Metaphor generation, Executive function, Figurative language comprehension

Acknowledgements

I would firstly like to thank my supervisor Dr Nancie Im-Bolter for all her help on this project. I am deeply grateful for her assistance and guidance over the course of the last two years. This project would not have happened without her and I feel very lucky to have had the chance to work with and learn from her.

I would also like to thank Dr Alba Agostino and Dr Marie Arsalidou for taking time to act as my committee members. Their feedback has been invaluable.

I finally I would also like to thank my cheerleaders: Kam, Sarah, Rebecca, and Christina, for their continued belief in me and endless patience and willingness to hear the oral version of this project from start to finish. Words cannot express my gratitude for their support throughout this process. In addition to thanking my friends and family, who I also could not have done this without.

Table of Contents

List of Tables

What's it Take to Get a Novel Metaphor Around Here? Exploring Potential Contributors to Novel Metaphor Production

Overview

Metaphors are a type of figurative, non-literal, language that make direct comparisons between unlike things (e.g., He is a rock; Skalicky, 2022). The simple nature of metaphors makes them particularly customizable to a speaker's communication needs; if no known phrase conveys the speaker's intended meaning, they only need to create a comparison to illustrate their point. Metaphors are frequently used to describe abstract concepts such as emotions and time, making them a crucial part of everyday communication and social relationships (Cappelli et al., 2018). Despite this there is limited research on the skills necessary for their production and research examining the production of novel metaphors is even more scarce. Novel metaphors refer to creative or entirely new phrases (e.g., moving at the speed of a biker going uphill) whereas conventional metaphors are well known or common phrases (e.g., moving at a snail's pace; Beaty & Silvia, 2013). The comprehension and use of metaphors begins in childhood and continues on into early adulthood (Levorato & Cacciari, 1995; Rundblad & Annaz, 2010; Van Herwegen et al., 2013), a time when executive function is also still maturing (Carriedo et al., 2016; Huizinga et al., 2006). Executive function is a higher order cognitive control process that is involved in goal-directed thoughts and behaviour (Friedman & Miyake, 2017; Miyake et al., 2000). Research shows a relation between metaphor production and executive function using self-report data (Menashe et al., 2020), but there is limited research directly investigating the role of executive function in this ability. Further, the majority of existing research on metaphor production allows other

forms of figurative language as well as metaphors (e.g., Beaty & Silvia, 2013; Menashe et al., 2020). The current research aims to clarify the role of executive function in the production of novel metaphors in early adulthood, a critical time for maturing social relationships that require more advanced figurative language skills (Nippold et al., 2017).

Metaphors

Metaphors are a ubiquitous component of language that make comparisons between two or more unlike things in order to convey meaning (Skalicky, 2022). This comparison is direct and is not explicitly signaled through the use of words such as "like" or "as". Being figurative, these comparisons are not intended to be understood literally. Instead, the comparison is meant to highlight a shared feature or attribute between the two components of a metaphor: the base (or vehicle) and the target (or topic; Bowdle & Gentner, 2005). The base of the metaphor is used to exemplify, or make salient, the feature of the target the speaker wants to highlight. For example, in the metaphor "The teacher is a tech wizard," *tech wizard* is the base that is being used to exemplify the characteristic of the *teacher*, which is the target of the metaphor. This metaphor communicates that the teacher is extremely proficient with technology and not that the teacher has wizard-like powers related to technology. However, this particular metaphor is an efficient way to suggest to the speaker that the teacher's abilities with technology may *seem* mystical. The comparison between the teacher and a tech wizard is concise and takes less time than explaining that the teacher knows a lot about technology and therefore can do things that seem magical or wizard-like.

This efficiency in communication is part of the reason for the widespread usage of metaphors. When well-constructed, despite their figurative or nonliteral nature,

metaphors are no more difficult to understand than literal language (Glucksberg, 1989). This means that in some cases they can be an even faster, more effective means of communication than literal words (Banaruee et al., 2019; Glucksberg, 2003; Thibodeau et al., 2019). Metaphors are also often better suited than literal language for communicating about topics which are otherwise hard to describe or explain such as emotions and abstract concepts like time (Lakoff, 1993). By comparing such topics to more concrete or tangible examples, metaphors are able to provide clear references to increase understanding. For example, when discussing time, an intangible concept, we rely almost entirely on metaphoric language (Lakoff, 1993). We talk about time "passing", which it cannot literally do since it is not a physical thing, yet we know this phrase means that we are progressing to a moment in the future point. We talk about "looking ahead" to the future even though we know that we cannot visually perceive future events that have not yet taken place. Similarly, when we talk about emotions, we can say that we are "seeing red" with anger or that our "skin crawls" from disgust, but we know that anger does not appear red in our vision and that disgust does not actually allow our skin to move in a crawling motion. Without metaphoric language it would be very difficult to communicate our intentions and feelings in such a vivid and efficient way.

We rely on metaphors to communicate about a nearly infinite number of abstract topics that are a part of our daily lives. This means that metaphors are not only a common part of our everyday language but are also engrained in the ways we conceptualize many topics whether we are aware of it or not (Lakoff, 1993). For example, we often use the concepts of motion and movement to represent progress towards a goal, using phrases such as "we're covering a lot of ground" or "we're at a standstill". We use the idea of a

journey as a representation for life, referring to expected achievements as *milestones*, large decisions as *crossroads,* and uncertainty about future plans as being *directionless*. It is nearly impossible to speak on topics like these without drawing from these broader conceptualizations, and therefore it is nearly impossible to communicate without using metaphor.

Given the extent to which metaphor is engrained in our language, efforts have been made to quantify its usage. Estimates for rates of metaphor usage vary based on the context in which they are examined (e.g., literature versus natural discourse, etc.) so reports of its frequency vary widely. Pollio et al. (1977), reported a rate of approximately six metaphors per minute of general speech whereas Steen (2011) reported that 13.6% of natural discourse was metaphoric. In a review of metaphor usage across different settings Dorst (2015) found that in academic settings metaphor usage was a high as 18.2% as compared to in conversations, where a rate of only 7.6% was observed. Despite the varying reports of metaphor frequency, it is clear that they make up a significant percentage of our language use. As such, the question of how we process such language has been an ongoing topic of interest in psychology.

Processing of Metaphors

The classic view of metaphor processing proposes that all language is first considered as literal, and if the literal interpretation cannot be determined, then a figurative meaning is considered (Lakoff, 1993). This type of processing would be both cognitively intensive and slow, which would contradict findings that metaphors are often comprehended at the same speed as literal statements (Glucksberg et al., 1997). Therefore, modern models of metaphor processing have taken different approaches to

explain how metaphors are comprehended. Some models propose straightforward processing mechanisms that rely on single cognitive strategies such as categorization (in which topics are placed into the conceptual categories that vehicles represent; Glucksberg, 2001) or comparison (in which vehicle and topic attributes are assessed and overlapping traits are applied to the topic; Clement & Gentner, 1991). For example, given the metaphor "time is money," a categorization strategy would place *time* in the same category as the vehicle *money* (*valuable* or *finite things*) and determine that the metaphor means that time is valuable or finite. Given the same metaphor, a comparison processing strategy would entail reviewing the salient attributes of both *money* (e.g., valuable, exchangeable, desirable, finite, etc.) and *time* (e.g., abstract, intangible, valuable, finite, etc.) and determining the meaning of the metaphor based on overlapping attributes (valuable, finite). However, it has been argued that a single mechanism cannot explain the processing of more complex or novel metaphors. As such, more nuanced theories, such as Bowdle and Gentner's (2005) Career of Metaphor (CoM) theory, incorporates a multifaceted approach to describing the processing of metaphors.

The CoM theory (Bowdle & Gentner, 2005) argues that previous models of metaphor processing have overlooked an important feature of metaphors, their type. Whether a metaphor is conventional or novel must be considered when theorizing about how they are processed. Conventional metaphors have an element of familiarity; they are typically well known and do not have to be processed as new phrases each time they are encountered. Novel or unfamiliar metaphors, which are heard for the first time when they are uttered cannot be processed in the same way. The CoM theory takes this into account and combines previous models to propose that both categorization and comparison are

involved in metaphor processing depending on the familiarity of the metaphors. Conventional metaphors are familiar enough to be processed via categorization, as conventional vehicles (e.g., the argument is garbage) are associated with conventional attributive categories (e.g., worthless; Bowdle & Gentner, 2005) and the mapping of this category to the topic is almost automatic. As a result, a comparison of the attributes between the vehicle and target is not needed.

The explanation provided by the CoM helps to account for comprehension time differences that have been observed between conventional and novel metaphors, wherein conventional metaphors are processed more quickly than novel metaphors (Blank, 1988; Gentner & Wolff, 1997). Categorization is a faster, more efficient form of processing that we can take advantage of when encountering conventional metaphors. Novel metaphors (e.g., the argument is war), however, must be processed via comparison. There is a lack of familiarity to aid in constructing a category for which the target will belong. As such, relevant attributes from the vehicle (e.g., strategy, attack) must be extracted and compared to the attributes of the topic, and through this comparison meaning is given to the metaphor. Although CoM theory was proposed with the goal of understanding metaphor comprehension, it serves as an important framework for understanding metaphor production, which relies on different but overlapping skills (Chiappe & Chiappe, 2007)

Conventional and Novel Metaphor Production

Given the prevalence of metaphors in communication, certain metaphors become engrained in culture resulting in some phrases taking on an additional, figurative meaning (Bowdle & Gentner, 2005). For example, a *gold mine* can refer to the literal hole in the

earth from which gold is extracted, or it can refer to something that is a source of value. Both meanings are equally well-understood to native speakers of English. If someone were to refer to a colleague as a "gold mine of ideas" it would be clear that what is intended from that statement is that the colleague in question is a source of good ideas. Conventional metaphors such as these employ comparisons that are familiar and commonly used, and rely on vehicles which are established within cultural knowledge and have highly salient meanings. Conventional metaphors can be contrasted to novel metaphors, which employ unique or creative vehicles to illustrate a point. For example, "the grape's a water balloon" is an original, or novel, metaphor which does not employ a previously established vehicle. However, and importantly, it is still comprehensible. The comparison indicates that the grape is bursting with juice, not that the grape is an actual latex balloon that is filled with water.

Both conventional and novel metaphors are used in communication (Lakoff, 1993; Steen, 2011), however, research on the frequency of their usage presents mixed findings. Lakoff (1993) posits that nearly all metaphors can be considered conventional because they are drawn from higher order conceptual mappings. For example, "life is a journey" is a higher order metaphor from which metaphors such as "I am spinning my wheels" or "I have hit a roadblock" would stem from. In support of this, Steen (2011) found that in adult's natural discourse, novel metaphors are used relatively infrequently compared to conventional metaphors. However, research by Barrow et al. (Barlow et al., 1977) appears to contradict Lakoff (1993) and Steen (2011). Barrow et al. found that both conventional and novel metaphors are used frequently by adults, with novel metaphors being more frequent than conventional ones. However, this research was conducted in a

psychotherapeutic context in which the topics of discussion would have been highly personal and likely involved emotions and other abstract concepts related to the self and each person's specific individual experiences. This context might explain the high frequency of novel metaphors, which would occur when the speaker is talking about unique and highly personal topics that requires novel language. Thus, the context and topic of conversation may contribute to the contradictory rates found for type of metaphor usage and further suggests the need to consider multiple factors in understanding metaphoric ability. This is in line with the CoM theory (Bowdler & Gentner, 2005), which implies that metaphor processing is a complex interaction between linguistic, cognitive, and contextual factors. Much of the research investigating the mechanisms underlying metaphor processing is focused on adults however, with less consideration regarding whether these mechanisms can be applied over development.

Metaphoric language production begins at a young age and continues to develop into young adulthood. Children as young as 2 years of age are capable of spontaneously producing simple metaphors (e.g., "your hair is grass" during play) in a manner that is both deliberate and appropriate for the given context (Billow, 1981). By 4 years children can produce novel metaphors when prompted (Read & Szokolszky, 2016). In general, the metaphors young children produce are basic and typically grounded in over-extensions (e.g., referring to an older man as "grandpa") or perceptual features (e.g., comparing hair to grass; Martín-González et al., 2024). The use of predominantly basic metaphors continues until age 9, after which more complex metaphors become more frequent (e.g., referring to internal states such as "boiling with anger"; Levorato & Cacciari, 2002).

Rates of metaphor use increase throughout childhood and into adolescence

(Levorato & Cacciari, 2002; M. R. Pollio & Pollio, 1974; Read & Szokolszky, 2016), but whether frequency of use plateaus in adolescence or young adulthood is not established. Some evidence suggests that there are significant differences in rates of metaphoric language use between early adolescence and young adulthood (Read & Szokolszky, 2016) and between mid-adolescence and young adulthood (Kasirer & Mashal, 2016). However, there is also research that shows no differences between young adults and adolescents aged 14 (Kasirer & Mashal, 2016) or 18 years (Levorato & Cacciari, 2002). Consideration of conventional versus novel metaphors does little to clarify this discrepancy. Kasirer and Mashal (2016) found that frequency of novel metaphor production did not differ between adolescents and young adults, whereas Read and Szokolzky (2016) found that it did. Despite these disagreements however, the current evidence points towards a trend in development in which metaphor usage increases up to adolescence and potentially into young adulthood.

Frequency of metaphor use may provide useful information regarding the trajectory of metaphoric production ability, but examining quality of metaphor use, as judged by aptness and appropriateness provides greater clarity. Levorato and Cacciari (2002) examined both frequency and quality of metaphor production in adolescents (aged 11 to 18 years old) and young adults (university students). They found that although adolescents and young adults were comparable with respect to quantity of metaphors produced, there was still a gap in the quality of these metaphors. Since the purpose of a metaphor is to better convey meaning, this discrepancy provides evidence that points towards the continued development of metaphoric ability beyond adolescence and into young adulthood. Given the complex nature of metaphors this makes sense; adolescents

may have mastered the basic form of metaphor and are able to use conventional metaphors at appropriate times. However, it is possible that they may have not yet mastered the ability to use metaphors creatively.

Recall that the CoM theory (Bowdle & Gentner, 2005) proposes that the processing of metaphors involves multiple factors and that understanding novel metaphors is more cognitively demanding than conventional metaphors. Similarly, the production of novel metaphors likely adds additional cognitive demands as an appropriate vehicle must be identified rather than just comprehended. The search for said vehicle must balance the consideration of all potential options and the overall communication goal, the intended meaning of the metaphor. Given the increased demands that producing a novel metaphor requires it is possible that the continued development of other cognitive skills involved in metaphoric ability may play a role in the mixed evidence found in the literature. The linguistic and cognitive skills required for conventional metaphor production may be comparable in adolescence and early adulthood, however the increased demands involved in novel metaphor production may require more advanced skills. One such skill that continues to develop through adolescence and into adulthood is executive function, which follows a similar developmental trajectory as metaphoric skills.

Executive Function

Executive function is a higher order cognitive control process that directs goal-oriented thoughts and behaviours (Miyake et al., 2000). Executive function comprises multiple distinct, yet related abilities, the most commonly postulated of which are shifting of mental sets (shifting), updating of working memory (updating), and inhibition of

prepotent responses (inhibition; Miyake et al., 2000; Miyake & Friedman, 2012). Shifting is the ability to purposefully move focused attention between different tasks (e.g., communicating the metaphoric goal of a slow moving car and searching one's memory for options such as *snail* or *molasses*), updating is the ability to monitor and modify information related to the task at hand so that new relevant information replaced old outdated information as needed, and inhibition is the ability to resist an initial automatic response in favour of a more appropriate goal-oriented action (Miyake et al., 2000).

Executive function develops rapidly throughout childhood, slows during adolescence, and peaks in young adulthood before stabilizing in adulthood and declining in older adulthood (Carriedo et al., 2016; Zelazo et al., 2004). However, within this general developmental pattern, the different components of EF appear to develop at different rates (Huizinga et al., 2006; Zelazo et al., 2004). Shifting develops quickly from early childhood, and seems to plateau or mature by adolescence (De Luca et al., 2003; Huizinga et al., 2006), while updating and inhibition continue to develop throughout this period into adulthood (Carriedo et al., 2016; De Luca et al., 2003). Despite the stability of EF throughout most of adulthood, young adulthood is a period in which EF is still developing, though at a lesser rate than previous stages (De Luca et al., 2003; Friedman et al., 2016; Huizinga et al., 2006). This developmental trajectory matches that of metaphoric ability, wherein both abilities have nearly reached maturity by the start of young adulthood but do not conclude development until later in this developmental stage (Levorato & Cacciari, 2002).

Executive Function and Novel Metaphor Production

Given the similar developmental trajectories between EF and metaphoric ability,

and the cognitive demands necessary for this ability as implied by the CoM theory (Bowdle & Gentner, 2005), it is reasonable to propose that EF contributes to novel metaphor production skill. As previously mentioned, the production of novel metaphors introduces increased cognitive demands as compared to conventional metaphors. These demands align well with the components encompassed within EF. For example, when considering the use of a metaphor during a conversational interaction (e.g., after a difficult exam: "That exam was a punch in the stomach"), the individual must search for an appropriate vehicle and update potential options, discarding previous but less effective options as new ones are found; shift between the memory search for a vehicle and the assessment of potential options with the communication goal in mind; and inhibit irrelevant, and in the case of novel metaphors, conventional options.

Though research on the production of novel metaphors is limited, there is some parallel literature that suggests a relation between EF and novel metaphor production. Correlations between novel metaphor production and skills related to EF such as fluid intelligence (Beaty & Silvia, 2013; Kasirer & Mashal, 2016), divergent thinking (Kasirer & Mashal, 2018), and phonemic fluency (Kasirer & Mashal, 2016) have all been used to suggest the possible involvement of EF in this ability. However, as this research does not directly examine core EF abilities, it cannot be taken as evidence of a relation between these metaphor production and EF on its own.

Support for the idea that EF is involved in novel metaphor production can be found through neuroimaging approaches. Studies investigating brain activity during novel metaphor production tasks show that regions associated with executive control processes are active during metaphor production tasks (Beaty et al., 2017; Benedek et al., 2014;

Chen et al., 2023). A comparison of brain activity during conventional versus novel metaphor production demonstrates that while there is overlapping activation during production of the two metaphor types, the production of novel metaphors resulted in the activation of a greater number of areas associated with EF (Chen et al., 2023). This points to the importance of EF in the production of novel metaphors in particular. This is in line with the CoM theory, which asserts that processing novel metaphors is more cognitively demanding than processing of conventional metaphors (Bowdle & Gentner, 2005). Neuroimaging studies provide more direct evidence for the involvement of EF in the production of novel metaphors. However, it is not clear whether EF in general or a specific component of EF contributes to this ability.

Menashe et al. (2020) offered a different perspective to research investigating EF and novel metaphor production by using a self-report measure of EF in daily living (Behavioral Rating Inventory of Executive function, Adult Version; BRIEF-A). They found a positive relation between overall EF score and novel metaphor production ability, as well as positive relations between specific EF skills such as the ability to initiate activities and organization skills and novel metaphor production. While these findings serve as direct evidence for the importance of EF in novel metaphor production, the use of a self-report measure of EF still leaves room for error in interpreting these results. As a self-report measure that asks about self-regulation in everyday tasks, the BRIEF-A is reliant on the individual's personal assessment of their behaviour (e.g., I find it hard to switch from one task to another) and has been criticized for not precisely measuring EF (Naglieri & Goldstein, 2014) and for exhibiting low convergent validity with behaviour-based EF measures (Toplak et al., 2013). Menashe et al. did use behavioural measures to

assess working memory and selective attention, two cognitive abilities which are related to updating and inhibition, respectively, according to Miyake et al.'s (2000) framework of EF. Working memory, or the ability to hold information in short term memory in order to perform some sort of operation (e.g., considering different meanings for the word "slow") is required for updating. Selective attention, or the ability to focus on the task at hand and ignore irrelevant information, requires the ability to inhibit the irrelevant information. Menashe et al. found that selective attention, but not working memory, contributed to novel metaphor performance, which suggests that inhibition may be involved. However, it is important to note that Menashe et al. allowed the use of similes (figurative expressions that use "like" or "as" to signal the comparison, e.g., "her hair is soft like silk") in their novel metaphor task making it unclear whether these findings reflect the role of EF in metaphor or simile production.

The Current Study

Menashe et al. (2020) are not the only researchers that included similes in their metaphor task performance score. In fact, nearly all studies examining metaphor production did the same, with some studies using prompts that actually encouraged the use of similes (e.g., "Eating that [food] was *like*…; Beaty & Silvia, 2013). In fact, only two studies (Beaty et al., 2017; Benedek et al., 2014) were strict about their definition of metaphor and did not allow participants to respond with similes. This is extremely problematic since similes are by definition, not metaphors, are used much earlier in development than metaphors, and most importantly, are less cognitively demanding to produce than metaphors (Bowdle & Gentner, 2005). In fact, this difference in cognitive demand appears to be the primary reason why some researchers chose to include similes

in their metaphor production studies, as demonstrated by Menashe et al. (2020) who stated that similes were encouraged to "facilitate the generation" of responses (p. 4). However, the inclusion of similes in metaphor research raises issues regarding the validity of the research findings regarding metaphor production in general, as well as the relation between EF and metaphor production ability. Given the prevalence of this methodological issue in the metaphor literature, there is still a large gap in our current understanding of the role of EF in novel metaphor production. Given this, the goal of the present study is to clarify our current understanding of the relation between novel metaphor production and EF in young adults through the use of an exclusively metaphor-based task alongside direct behavioural measures of EF derived from Miyake et al.'s (2000) framework. Further, the present study also examines whether there are any differences between young adults who are able to produce novel metaphors and those who cannot.

Hypotheses

1. Based on the findings that a self-report measure of EF and a behavioral measure of selective attention were positively related to novel metaphor production (Menashe et al., 2020), and that overall figurative language skill and metaphor comprehension rely on overlapping abilities (Benedek et al., 2014), individual differences in EF and figurative language comprehension are expected to predict novel metaphor production.
2. Consistent with the first hypothesis, it is expected that EF skills (updating, shifting, and inhibition) and figurative language comprehension will differ between individuals who can create novel metaphors compared to those who are

not able to create novel metaphors.

Method

Participants

Undergraduate students from Trent University were recruited using the online research system SONA as part of a larger study investigating social cognition in young adults. Ethical approval was granted by the Trent University Research Ethics Board. Participants received partial course credit in an introductory psychology course in return for their participation. The study sample consisted of 112 young adults (80.36% Female, 82.14% English as a first language) aged 17 to 29 years (M = 21.06, SD = 2.89). All participants provided written consent before taking part in the study.

Procedure

Participants took part in an individualized testing session which typically lasted between 2.5 to 3 hours. During this session participants completed a background questionnaire and tasks that measured nonverbal intelligence, figurative language comprehension, metaphor production, and EF (updating, shifting, and inhibition), as well as other measures that were not relevant to the current study. Task order for each session was randomly assigned with the background questionnaire always being administered first. Breaks were provided during the testing session as needed.

Measures

Background Information

Background information about each participant was obtained via a self-report questionnaire. Participants reported their first language, and the age at which they began learning English if it was not their first language. Socioeconomic status (SES) was

estimated based on parental occupations using the Blishen Scale (Blishen et al., 1987). The higher of the two parental SES estimates was used for analyses.

Estimated Non-Verbal Intelligence

The *Matrix Reasoning* subtest of the *Wechsler Abbreviated Scale of Intelligence—Second Edition* (WASI-II; Wechsler, 2011) was used to estimate non-verbal intelligence. This task is comprised of 30 items consisting of a 2 x 2 matrix of a visual pattern in which one block of the matrix was blank. Five options for completing the visual pattern were provided below each item and participants were instructed to choose the option that would best complete the pattern. Standardized scores were derived from total raw scores and used for all analyses.

Figurative Language

Figurative language comprehension and production was assessed with two measures.

The *Figurative Language* subtest of the *Clinical Evaluation of Language Fundamentals-Fifth Edition, Metalinguistics* was used to assess figurative language comprehension. Participants were presented with different idioms, which are nonliteral phrases such as "it's raining cats and dogs", within a short context and asked to describe the meaning of the idiom. They are then asked to identify the correct meaning of the idiom from four options. Raw scores were converted to standardized scores which were then used for all analyses.

The *novel metaphor production task*, adapted from Silvia and Beaty (2021), was used to measure the ability to create novel metaphors. After being provided with the definition of a metaphor and multiple examples of simple (e.g., "life is a highway,"

Cochran, 1991) and complex metaphors (e.g., “our words are but crumbs that fall down from the feast of the mind,” Gibran, 1946, p. 23), participants were instructed to come up a metaphors based on three different prompts that asked them to describe: “the most boring class you’ve been in”, “the worst movie you’ve seen”, and “the most disgusting food you’ve eaten”. Participants were asked to create two metaphors per prompt for a total of six metaphors. Participants were specifically instructed not to use the words “like” or “as” in their metaphors and to create metaphors that were as creative, humorous, original, or interesting as they could think of. Emphasizing the importance of generating a novel or creative prompt ensures that participants understand the purpose of the task and is common practice for measurements of creativity (Nusbaum et al., 2014). All participants completed this task by writing their responses on paper and were given no time limit. Participants’ responses were reviewed to ensure they were full sentences. If the written responses were not full sentences (e.g., “older than the pharaohs in the tombs”) then participants were asked to complete them.

All responses on the metaphor task were rated by two coders to first determine if the response was metaphorical. Reliability for metaphor determination was high for prompts one ($\kappa = .96$, $p < .0001$), two ($\kappa = .98$, $p < .001$) and three ($\kappa = .98$, $p < .0001$). The two coders then rated the novelty of each metaphorical response on a three-point scale (1-3) with a score of 1 for conventional metaphors (e.g., stock phrases, “this movie is trash”); a score of 2 for metaphors that were modified or derived from a conventional metaphor (e.g., “this movie is all the waste of the world dumped onto the screen”); and a score of 3 for novel metaphors (e.g., “*The Vampire Diaries* was made on a dairy farm that specializes in cheese”). Reliability for novelty was high for prompts one ($\kappa = .91$, $p <$

.001), two ($\kappa = .98$, $p < .001$), and three ($\kappa = .96$, $p < .001$). Disagreements for coding responses as metaphors and for novelty were resolved through discussion until a consensus was reached.

Mean novelty scores for each participant were calculated based on the novelty ratings of all six responses. Additionally, participants were put into one of four groups: 1) *no metaphors*, participants who did not provide any metaphorical responses (e.g., sentence fragments, literal statements, or other forms of figurative language such as similes or hyperboles), 2) *conventional*, participants who provided only conventional metaphorical responses, 3) *modified*, participants who provided at least one modified conventional response but no novel metaphors, and 4) *novel*, participants who provided at least one novel response throughout the task.

Executive Function

Based on Miyake et al. (2000), three components of executive function were assessed: shifting of mental sets (shifting), updating of the contents of working memory (updating), and inhibition of prepotent responses (inhibition).

Shifting was measured with the *plus-minus task* (Miyake et al., 2000). This task consists of three columns of 30 numbers each (ranging from 1-99) presented on an 8 x 11 ½ inch sheet of paper. Participants were required to add 3 to each number in the first column; subtract 3 from each number in the second column; and switch between adding and subtracting 3 from each number in the third column, beginning with addition. Participants were timed and asked to write their answers on the sheet as quickly as possible, but to slow down if necessary to reduce errors. The time to complete each column was hand recorded with a digital stopwatch.

A ratio score was calculated, based on the mean time to complete columns one and two and the time to complete column three ([Column 3 – mean (Columns 1 and 2)]/ mean (Columns 1 and 2)). Lower ratio scores represented better shifting abilities. These ratio scores were used for all analyses.

The *letter memory task* (Miyake et al., 2000) was used to assess updating. The letter memory task is a computerized task where participants are presented with a series of single upper case consonant letters presented at 2 second intervals. Participants were instructed to say out loud the most recent four letters in the order that they appeared. For example, if the letter sequence B K C D Z was presented, participants would say B… BK… BKC…. BKCD…KCDZ. The task consisted of four practice trials (with 5, 7, 9, and 11 letters) and 16 test trials that ranged in length from 5-11 letters. The updating score represented the proportion of letters that were successfully updated. Therefore, higher proportion scores indicate better updating.

Inhibition was assessed with the *Stroop task* (Stroop, 1935). This task is comprised of two conditions: neutral and Stroop. The neutral condition was comprised of stimuli that were between three and six upper case Xs printed in different ink colours (e.g., red, green). The Stroop condition consisted of colour words (e.g., red, green) printed in different ink colours (e.g., the word *red* printed in green ink). In each condition, stimuli were arranged in three columns of ten on an 8 x 11 ½ sheet of paper. There were three sheets for each condition for a total of six trials. The presentation of these stimuli was counterbalanced such that no more than two trials in a row were the same condition. Automaticity of reading and colour vision was assessed prior to beginning the Stroop task. Participants were instructed to name the ink colour of every item presented on each

sheet of paper as quickly as possible. A practice trial was provided to ensure participants understood the task before the test trials were administered. The time to complete each trial was hand recorded using a digital stopwatch. Errors were also recorded. A ratio score was calculated based on the mean time to complete the neutral condition and the Stroop condition ([Stroop – Neutral]/Neutral; Barzykowski et al., 2021). Lower mean scores indicated better inhibitory control. These ratio scores were used for all analyses.

Results

Data Screening

Prior to analysis data was screened. Participants who did not complete all measures were removed (n = 4). Data were examined for normality and outliers, and assumptions were assessed. Distributions of all measures were deemed normal except metaphor scores, which were positively skewed. This is expected for developmental skills, which have a stage like progression where proficiency is expected for the lower stage (i.e., conventional metaphors) and developing proficiency is expected for the highest stage (i.e., novel metaphors; (Locke, 1994). For the plus-minus and Stroop tasks, z-scores were calculated based on mean and standard deviation for total task errors and then used to identify problematic cases. High errors on these tasks impact their reliability; a large number of errors on the plus-minus task suggest that shifting between adding and subtracting may not actually be occurring and on the Stroop task suggest a failure to inhibit the automatic response of reading. As such, cases with z-scores greater than three were examined. This resulted in the removal of five participants, four with errors ranging from 19-26 on the plus-minus task and one with 36 errors on the Stroop task. Finally, outliers were assessed using boxplots, with scores falling more than three standard

deviations from the mean being categorized as such. No outliers were detected through this assessment. After screening, 112 participants remained for analysis. Following this, necessary assumptions for the planned ANCOVAs were assessed. Homogeneity of variance of all variables was assessed using Levene's test and found to be satisfied (all tests $p > .287$). Homogeneity of regression slopes was assessed and all found to be satisfied (all interactions $p > .132$). All other assumptions (e.g., independence of variables, normality, etc.) were also found to be satisfied.

Background Characteristics

Overall, participants had a mean age of 21.06 years ($SD = 2.89$, range 17-29 years), the majority of whom were female (80.36%, $n = 90$), came from middle class backgrounds (SES $M = 55.52$, $SD = 15.03$, range 25.53-101.32), and had an estimated nonverbal IQ that was within the average range ($M = 10.42$, $SD = 2.72$, range: 3-18). The majority of participants spoke English as their first language (82.14%, $n = 92$). Of those who acquired English as an additional language, the mean age of acquisition was 4.82 years ($SD = 3.21$, range 1-15 years).

As previously described, performance on the *Novel Metaphor Task* was used to place participants into one of four groups: 27.68% (31) of participants fell in the *No Metaphors* group, 46.43% (52) of participants fell in the *Conventional Metaphors* group, 15.81% (17) of participants fell in the *Modified Conventional* group, and 10.71% (12) of participants fell in the *Novel Metaphors* group. There were no group differences with respect to English as a first language status (X^2 (3, $N = 112$) = 0.97, $p = .809$; see Table 1 for frequencies).

Table 1

Background Characteristics by Metaphor Group

	No Metaphor $n = 31$		Conventional Metaphor $n = 52$		Modified Conventional $n = 17$		Novel Metaphor $n = 12$		F / X^2
% Female (*n*)	83.87 (26)		80.77 (42)		76.47 (13)		75.00 (9)		0.63
% English First Language (*n*)	87.10 (27)		80.77 (42)		76.47 (13)		83.33 (10)		0.87
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	
Age	21.81	3.08	20.77	2.59	20.35	2.85	21.42	3.55	1.27
SES	57.00	18.78	53.44	13.09	53.67	12.45	62.85	13.83	1.47
Estimated Nonverbal IQ	9.94	2.21	8.44	1.97	11.65	2.85	11.83	1.95	3.20*

Note. *$p < .05$.

Group differences in age, SES, and estimated nonverbal IQ were assessed using one-way analysis of variance (ANOVA) and Tukey's Honestly Significant Difference (HSD) tests. Groups did not differ with respect to age or SES ($\eta^2 = .04$; see Table 1 for means and standard deviations), but there was a significant difference between groups with respect to estimated nonverbal IQ ($\eta^2 = .08$; see Table 1 for means and standard deviations). Post hoc analysis using Tukey's HSD revealed no significant differences between the groups (all *p*s > .1). Visual examination of the means suggested a difference between groups that could create some sort of a novel metaphor (*Modified Conventional* and *Novel* groups) and the groups that could not (*No Metaphors* and *Conventional Metaphor*). To explore this idea, an additional ANOVA was conducted with those two groups (*Not Novel* vs. *Novel*). This analysis revealed significantly higher nonverbal IQ in the new Novel group in comparison to the Not Novel group (see Table 2 for summary of other background characteristics). As such all of the following analyses included estimated nonverbal IQ as a covariate in group analyses.

Table 2

Background Characteristics by Not Novel and Novel Metaphor Groups

	Not Novel $n = 83$		Novel $n = 29$		F / X^2
% Female (*n*)	80.40 (68)		19.60 (22)		.48
% English First Language (*n*)	17.90 (69)		82.10 (23)		.64
	M	*SD*	*M*	*SD*	
Age	21.16	2.81	20.79	3.14	0.34
SES	54.78	15.47	57.75	13.64	0.79
Estimated Nonverbal IQ	9.96	2.66	11.72	2.48	9.73**

Note. $^{**}p < .01$.

Individual Differences in Novel Metaphor Production Ability

To test the first hypothesis regarding individual differences in novel metaphor production, Pearson Product Moment correlations were conducted. First, background variables were correlated with figurative language, novel metaphor production, and executive function (shifting, updating, inhibition). Significant correlations were found between estimated nonverbal intelligence and updating ($r = .32$, $p < .001$) and between SES and updating ($r = .27$, $p = .004$). As such, both estimated nonverbal intelligence and SES were partialled out of the remaining correlations (see Table 3). There were no significant correlations between any of the measures, indicating no relation between figurative language and novel metaphor production ability, and executive function.

Group Differences in Novel Metaphor Production Ability

To test the second hypothesis, regarding group differences in novel metaphor production, analyses were conducted using analysis of covariance (ANCOVA) with estimated nonverbal intelligence as a covariate.

Table 3

Means, Standard Deviations, and Partial Correlations Controlling for Estimated Nonverbal Intelligence and Socioeconomic Status

	Variable	*M*	*SD*	1	2	3	4
1	Figurative Language Comprehension	8.44	1.84				
2	Novel Metaphor Production	0.48	.51	.23			
3	Shifting	.32	.23	-.02	.02		
4	Updating	.32	.19	.17	.24	-.15	
5	Inhibition	.46	.16	-.02	-.19	.04	-.14

Note: Figurative Language Comprehension is represented by a standard scaled score; Novel Metaphor Production is represented by a mean novelty score; Shifting is represented by a ratio score; Updating is represented by a proportion correct score; and Inhibition is represented by a ratio score.

Figurative Language Comprehension

Analyses revealed no group differences with respect to figurative language comprehension whether analyses were conducted with the original four groups ($F(3,107) = 1.09$, $p = .356$, $\eta^2 = .03$) or with the two new groups ($F(1,110) = 0.02$, $p = .891$, $\eta^2 = .00$).

Executive Function

Analyses revealed no group differences in shifting ($F(3,107) = 0.11$, $p = .953$, $\eta^2 = .00$; $F(2,112) = 0.03$, $p = .859$, $\eta^2 = .00$), updating ($F(3,107) = 0.30$, $p = .825$, $\eta^2 = .01$; $F(2,112) = 0.19$, $p = .665$, $\eta^2 = .00$), or inhibition ($F(3,107) = 0.14$, $p = .933$, $\eta^2 = .00$; $F(2,112) = 0.00$, $p = .993$, $\eta^2 = .00$) with respect to the original four groups or the two new groups, respectively.

Discussion

The current study examined the relations between EF and figurative language comprehension with novel metaphor production in young adults. Neither hypothesis

regarding the contribution of figurative language comprehension nor EF to individual or group differences in novel metaphor production was supported. However, estimated nonverbal intelligence did differ across the metaphor type groups.

Figurative Language Comprehension and Novel Metaphor Production

Individual differences in figurative language comprehension do not significantly predict novel metaphor production in the current study, which appears contrary to previous research that shows that they rely on overlapping abilities (Benedek et al., 2014; Chiappe & Chiappe, 2007). It seems unlikely that one could create figurative statements, such as metaphors, without an ability to understand them. It is possible however, that figurative language comprehension, acquired in childhood (Martín-González et al., 2024), is a foundational ability that is necessary, but not sufficient for advanced figurative language skills such as creating novel metaphors. This could explain the lack of relation between the two in the current study. Further research examining figurative language comprehension and novel metaphor production at different developmental stages may help to clarify the relation between the two.

Executive Function and Novel Metaphor Production

Contrary to previous research (e.g., Beaty et al., 2017; Benedek et al., 2014; Menashe et al., 2020), we find no relation between any of the components of EF and novel metaphor production. One reason for this may be the methodology used in the present study, which is only one of three studies, to our knowledge, that excluded similes when considering novel metaphor production. However, the other two studies that collected exclusively metaphoric responses (Beaty et al., 2017; Benedek et al. 2014) scaffolded novel metaphor production in their participants by instructing them to replace

an adjective with a metaphoric vehicle (Benedek et al., 2014). For example, the participant would be given the phrase, “the lamp is *glaring*”, and asked to replace the word “glaring”. This type of task provides participants with a direct comparison and ensures that the final product is a metaphor. This serves to decrease the task demand since providing a stem removes the need to construct the comparison itself, and instead requires a search for an appropriate vehicle only. As a result, participant performance, which is based on choosing a more or less creative vehicle can vary. The task used in the present study required participants to structurally create a full novel metaphor after being given a general topic (e.g., food), which may have resulted in poor performance due to the higher task demands in comparison to previous research (i.e., a floor effect).

Another factor that may have contributed to the present findings, and is related to the idea of higher task demands, is the research context of the current study. Previous literature indicates that rates of metaphor usage may be related to social context (Barlow et al., 1977; Steen et al., 2010). For example, contexts that require explanation of difficult concepts, such as in academic lectures, tend to exhibit higher rates of metaphor usage in comparison to contexts such as storytelling (Steen et al., 2010). Further, contexts with high emotional investment in which personal topics are addressed, such as therapy, tend to elicit more novel than conventional metaphors (Barlow et al., 1977). The present study required participants to create novel metaphors on demand in a structured lab setting. While such settings are useful for ensuring specific types of language are elicited, they can make it more difficult for individuals to spontaneously express certain language forms compared to more natural, social contexts where the individual has a personal or emotional investment in the conversational topic (Dunbar, 2001).

Although the novel metaphor task in the present study utilized prompts that had high topic salience (e.g., personal experiences related to negative emotions; Silvia & Beaty, 2021), it is possible that participants were not emotionally invested in the task. As a result, they may have struggled to find an appropriate topic and vehicle to create a novel metaphor. Emotional investment may be an important factor for the production of novel metaphors, as demonstrated by the relatively high rates of novel metaphors used in therapeutic contexts (Barlow et al., 1977). In this context, people are typically explaining their emotions and experiences in a highly personalized discussion they are emotionally invested in. Emotional investment, and the accompanying importance of being able to fully express one's communicative intentions may therefore be important for the production of novel metaphors. This possibility would also explain the high rates of metaphor usage observed in academic settings (Steen et al., 2010), and specifically in science settings (McGrath & Liardét, 2023), wherein speakers are invested in explaining theories and findings, which may be difficult to explain directly. Future research is needed to explore how personal or emotional investment might be related the creation of novel metaphors. For example, presenting the metaphor task as a semi-structured interview on a topic of particular importance to the participant might allow for more natural conversation and aid in increasing topic salience and emotional investment in the communication.

In the present study only ~11% of the young adult participants were able to create novel metaphors. Equally surprising is that ~28% of the young adult participants in the current study were not able to create any metaphors at all. This poor performance was despite the provision of numerous examples of metaphors and coaching from the

examiners to help the participants create a metaphor. This floor effect is itself interesting, as the current study sample is composed of university students from primarily middle-class backgrounds. This would usually be considered a limitation of the study, but in this case provides an important perspective to our findings, which are contrary to research that demonstrates a positive association between education and socioeconomic status with language performance (Ferreira-Correia et al., 2024; Gathercole et al., 2016). Considering why this association did not hold in our findings may shed light on metaphor production ability in particular. One explanation for the low metaphoric ability in the current study is the COVID-19 pandemic, which resulted in school closures and a significant restriction in social interactions. Research examining school closures and learning outcomes in children in grades 3-8 show negative correlations between literacy achievement and length of school closures, wherein longer school closures resulted in worse literacy scores post pandemic (Kennedy & Strietholt, 2023; Kuhfeld et al., 2023; Molnár & Hermann, 2023; Patrinos, 2023). Research is needed that specifically examines level of education, socioeconomic status, and metaphor production to clarify these associations. Further, future research recruiting young adults whose education was impacted by the pandemic should assess vocabulary and print exposure to account for the changes in literacy and education that resulted from the pandemic.

In sum, while poor performance is not atypical in research examining novel metaphor production (e.g., Silvia & Beaty, 2021), the performance in the present study did result in a floor effect. This floor effect suggests that creating novel metaphors may be inherently difficult in ways that EF skill, which is used to direct and control problem solving, does not address.

Nonverbal IQ and Novel Metaphor Production

In the current study, nonverbal IQ predicts novel metaphor production such that mean nonverbal IQ scores are significantly different between groups that can create a novel metaphor and those that cannot. This indicates that nonverbal IQ may play a role in novel metaphor production, which is a language-based task. Although surprising, this is not entirely unprecedented. For example, Kasirer and Mashal (2016) also found a significant relation between nonverbal IQ and novel metaphor production. It is possible that general problem solving and abstract reasoning represented by nonverbal IQ is important for the creative aspect of novel metaphor production (e.g., ability to identify underlying commonalities between the topic and vehicle). Identifying an appropriate vehicle requires comparing the attribute of the topic that one intends to communicate against the salient attributes of potential vehicles in order to find the best exemplar of this attribute. Abstract problem solving skills could facilitate the vehicle search process by aiding in the identification and assessments of potential vehicles and therefore enhance better novel metaphor production ability.

Support for this idea can be found in research that shows a significant association between tasks that measure fluid intelligence (also known as abstract reasoning) and novel metaphor production (Beaty & Silvia, 2013; de Barros et al., 2010; Silvia & Beaty, 2012), but not conventional metaphor production (Beaty & Silvia, 2013; Silvia & Beaty, 2012). Abstract reasoning abilities may be important in supporting the search and evaluation of potential vehicles for novel metaphors. This possibility is supported by the association of nonverbal IQ with other measures of creativity, such as higher creative output (Benedek et al., 2012). Future research should continue to investigate the relation

between nonverbal IQ and novel metaphor production using other measures of nonverbal IQ, such as analogical reasoning tasks, to determine how specific nonverbal abilities contribute to novel metaphor production.

Conclusion

In conclusion, the present study shows no relation between figurative language comprehension or EF and the ability to create novel metaphors. Methodological differences might help to explain the inconsistencies between our findings and those of previous research. It is noteworthy however, that the young adults in the current study found it particularly difficult to create novel metaphors and that nonverbal IQ appears to play a role in metaphoric ability. Research that addresses developmental stages, social contextual factors, education, socioeconomic status, and abstract reasoning skills are needed to further our understanding of the factors involved in novel metaphoric ability.

References

Banaruee, H., Khoshsima, H., Zare-Behtash, E., & Yarahmadzehi, N. (2019). Several reasons behind using metaphor: A cognitive perspective on metaphoric language. *NeuroQuantology. 17*(3), 108-113. http://doi.org/10.14704/nq.2019.17.3.2001

Barlow, J. M., Pollio, H. R., & Fine, H. J. (1977). Insight and figurative language in psychotherapy. *Psychotherapy*, *14*(3), 212–222. https://doi.org/10.1037/h0086530

Barzykowski, K. J., Wereszczyński, M., Hajdas, S., & Radel, R. (2021). An Inquisit-web protocol for calculating composite inhibitory control capacity score: An individual differences approach. *MethodsX*, *8*. https://doi.org/10.1016/j.mex.2021.101530

Beaty, R. E., & Silvia, P. J. (2013). Metaphorically speaking: Cognitive abilities and the production of figurative language. *Memory & Cognition*, *41*(2), 255–267. https://doi.org/10.3758/s13421-012-0258-5

Beaty, R. E., Silvia, P. J., & Benedek, M. (2017). Brain networks underlying novel metaphor production. *Brain and Cognition*, *111*, 163–170. https://doi.org/10.1016/j.bandc.2016.12.004

Benedek, M., Franz, F., Heene, M., & Neubauer, A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences*, *53*(4), 480–485. https://doi.org/10.1016/j.paid.2012.04.014

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, *46*, 73–83. https://doi.org/10.1016/j.intell.2014.05.007

Billow, R. M. (1981). Observing spontaneous metaphor in children. *Journal of*

Experimental Child Psychology, *31*(3), 430–445. https://doi.org/10.1016/0022-0965(81)90028-X

Blank, G. D. (1988). Metaphors in the Lexicon. *Metaphor and Symbolic Activity*, *3*(3), 21–36. https://doi.org/10.1207/s15327868ms0301_2

Blishen, B. R., Carroll, W. K., & Moore, C. (1987). The 1981 socioeconomic index for occupations in Canada. *Canadian Review of Sociology/Revue Canadienne de Sociologie*, *24*(4), 465–488. https://doi.org/10.1111/j.1755-618X.1987.tb00639.x

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216. https://doi.org/10.1037/0033-295X.112.1.193

Cappelli, G., Noccetti, S., Arcara, G., & Bambini, V. (2018). Pragmatic competence and its relationship with the linguistic and cognitive profile of young adults with dyslexia. *Dyslexia*, *24*(3), 294–306. https://doi.org/10.1002/dys.1588

Carriedo, N., Corral, A., Montoro, P. R., Herrero, L., & Rucián, M. (2016). Development of the updating executive function: From 7-year-olds to young adults. *Developmental Psychology*, *52*(4), 666–678. https://doi.org/10.1037/dev0000091

Chen, Q., He, R., Sun, J., Ding, K., Wang, X., He, L., Zhuang, K., Lloyd-Cox, J., & Qiu, J. (2023). Common brain activation and connectivity patterns supporting the generation of creative uses and creative metaphors. *Neuropsychologia*, *181*, 108487. https://doi.org/10.1016/j.neuropsychologia.2023.108487

Chiappe, D. L., & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of Memory and Language*, *56*(2), 172–188. https://doi.org/10.1016/j.jml.2006.11.006

Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in

analogical mapping. *Cognitive Science*, *15*(1), 89–132. https://doi.org/10.1207/s15516709cog1501_3

Cochran, T. (1991, September 20). *Life is Highway* [Song]. Capitol.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the CANTAB. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, *25*(2), 242–254. https://doi.org/10.1076/jcen.25.2.242.13639

Dorst, A. G. (2015). More or different metaphors in fiction? A quantitative cross-register comparison. *Language and Literature: International Journal of Stylistics*, *24*(1), 3–22. https://doi.org/10.1177/0963947014560486

Dunbar, K. (2001). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *Analogy: Perspectives from cognitive science* (pp. 313–334). The MIT Press. https://doi.org/10.7551/mitpress/1251.003.0013

Ferreira-Correia, A., Banjo, H., & Israel, N. (2024). Phonemic verbal fluency in non-WEIRD populations: Demographic differences in performance in the Controlled Oral Word Association Test-FAS. *African Journal of Psychological Assessment*, *6*, 152. https://doi.org/10.4102/ajopa.v6i0.152

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late

adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, *52*(2), 326–340. https://doi.org/10.1037/dev0000075

Gathercole, V. C. M., Kennedy, I., & Thomas, E. M. (2016). Socioeconomic level and bilinguals' performance on language and cognitive measures. *Bilingualism: Language and Cognition*, *19*(5), 1057–1078. https://doi.org/10.1017/s1366728915000504

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language*, *37*(3), 331–355. https://doi.org/10.1006/jmla.1997.2527

Gibran, K. (1946). *Sand and Foam*. Alfred A. Knopf, Inc.. (Original work published 1926)

Glucksberg, S. (1989). Metaphors in conversation: How are they understood? Why are they used? *Metaphor and Symbol*, *4*(3), 125–143. https://doi.org/10.1207/s15327868ms0403_2

Glucksberg, S. (2001). Beyond comparison property attribution. In S. Glucksberg (Ed.), *Understanding figurative language: From metaphor to idioms*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195111095.003.0004

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, *7*(2), 92–96. https://doi.org/10.1016/S1364-6613(02)00040-2

Glucksberg, S., McGlone, M. S., & Manfredi, D. (1997). Property attribution in metaphor comprehension. *Journal of Memory and Language*, *36*(1), 50–67. https://doi.org/10.1006/jmla.1996.2479

Huizinga, M., Dolan, C. V., & Van Der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis.

Neuropsychologia, *44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *14, 89*, 81.

Kasirer, A., & Mashal, N. (2016). Comprehension and generation of metaphoric language in children, adolescents, and adults with dyslexia. *Dyslexia*, *23*(2), 99–118. https://doi.org/10.1002/dys.1550

Kasirer, A., & Mashal, N. (2018). Fluency or similarities? Cognitive abilities that contribute to creative metaphor generation. *Creativity Research Journal*, *30*(2), 205–211. https://doi.org/10.1080/10400419.2018.1446747

Kennedy, A. I., & Strietholt, R. (2023). School closure policies and student reading achievement: Evidence across countries. *Educational Assessment, Evaluation and Accountability*, *35*(4), 475–501. https://doi.org/10.1007/s11092-023-09415-4

Kuhfeld, M., Lewis, K., & Peltier, T. (2023). Reading achievement declines during the COVID-19 pandemic: Evidence from 5 million U.S. students in grades 3–8. *Reading and Writing*, *36*(2), 245–261. https://doi.org/10.1007/s11145-022-10345-8

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and Thought* (2nd ed., pp. 202–251). Cambridge University Press. https://doi.org/10.1017/CBO9781139173865.013

Levorato, M. C., & Cacciari, C. (1995). The effects of different tasks on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology*, *60*, 261–283. https://doi.org/10.1006/jecp.1995.1041

Levorato, M. C., & Cacciari, C. (2002). The creation of new figurative expressions:

Psycholinguistic evidence in Italian children, adolescents and adults. *Journal of Child Language*, *29*(1), 127–150. https://doi.org/10.1017/S0305000901004950

Locke, J. L. (1994). Phases in the child's development of language. *American Scientist*, *82*(5), 436–445.

Martín-González, I., R. Ronderos, C., Castroviejo, E., Schroeder, K., Lossius-Falkum, I., & Vicente, A. (2024). That kid is a grasshopper! Metaphor development from 3 to 9 years of age. *Journal of Child Language, 52*(4), 945-970. https://doi.org/10.1017/S0305000924000187

McGrath, D., & Liardét, C. (2023). Grammatical metaphor across disciplines: Variation, frequency, and dispersion. *English for Specific Purposes*, *69*, 33–47. https://doi.org/10.1016/j.esp.2022.09.003

Menashe, S., Leshem, R., Heruti, V., Kasirer, A., Yair, T., & Mashal, N. (2020). Elucidating the role of selective attention, divergent thinking, language abilities, and executive functions in metaphor generation. *Neuropsychologia*, *142*, 107458. https://doi.org/10.1016/j.neuropsychologia.2020.107458

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Molnár, G., & Hermann, Z. (2023). Short- and long-term effects of COVID-related kindergarten and school closures on first- to eighth-grade students' school readiness skills and mathematics, reading and science learning. *Learning and Instruction*, *83*, 101706. https://doi.org/10.1016/j.learninstruc.2022.101706

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*(2), 111–121. https://doi.org/10.1111/j.2044-8295.1990.tb02349.x

Naglieri, J. A., & Goldstein, S. (2014). Assessment of executive function using rating scales: Psychometric considerations. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 159–170). Springer. https://doi.org/10.1007/978-1-4614-8106-5_10

Nippold, M. A., Frantz-Kaspar, M. W., & Vigeland, L. M. (2017). Spoken language production in young adults: Examining syntactic complexity. *Journal of Speech, Language, and Hearing Research*, *60*(5), 1339–1347. https://doi.org/10.1044/2016_JSLHR-L-16-0124

Nusbaum, E. C., Silvia, P. J., & Beaty, R. E. (2014). Ready, set, create: What instructing people to "be creative" reveals about the meaning and mechanisms of divergent thinking. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(4), 423–432. https://doi.org/10.1037/a0036549

Patrinos, H. A. (2023). The longer students were out of school, the less they learned. *Journal of School Choice, 17*(2), 161–175. https://doi.org/10.1080/15582159.2023.2210941

Pereira de Barros, D. P., Primi, R., Miguel, F. K., Almeida, L. S., & Oliveira, E. P. (2010).

Metaphor creation: A measure of creativity or intelligence? *European Journal of Education and Psychology, 3*(1), 103-115.

Pollio, H. R., Barlow, J. M., Fine, H. J., & Pollio, M. R. (1977). *Psychology and the Poetics of Growth: Figurative Language in Psychology, Psychotherapy, and Education*. Routledge. https://doi.org/10.4324/9781003454656

Pollio, M. R., & Pollio, H. R. (1974). The development of figurative language in children. *Journal of Psycholinguistic Research*, *3*(3), 185–201. https://doi.org/10.1007/BF01069237

Read, C., & Szokolszky, A. (2016). A developmental ecological study of novel metaphoric language use. *Language Sciences*, *53*, 86–98. https://doi.org/10.1016/j.langsci.2015.07.003

Rundblad, G., & Annaz, D. (2010). Development of metaphor and metonymy comprehension: Receptive vocabulary and conceptual knowledge. *British Journal of Developmental Psychology*, *28*(3), 547–563. https://doi.org/10.1348/026151009X454373

Silvia, P. J., & Beaty, R. E. (2012). Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence*, *40*(4), 343–351. https://doi.org/10.1016/j.intell.2012.02.005

Silvia, P. J., & Beaty, R. E. (2021). When figurative language goes off the rails and under the bus: Fluid intelligence, openness to experience, and the production of poor metaphors. *Journal of Intelligence*, *9*(1), 2. https://doi.org/10.3390/jintelligence9010002

Skalicky, S. (2022). Liquid gold down the drain: Measuring perceptions of creativity

associated with figurative language and play. *Cognitive Semantics*, *8*(1), 79–108. https://doi.org/10.1163/23526416-bja10018

Steen, G. J. (2011). The contemporary theory of metaphor—Now new and improved! *Review of Cognitive Linguistics*, *9*(1), 26–64. https://doi.org/10.1075/ml.9.1.03ste

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., & Krennmayr, T. (2010). Metaphor in usage. *Cognitive Linguistics*, *21*(4), 765–796. https://doi.org/10.1515/cogl.2010.024

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Thibodeau, P. H., Matlock, T., & Flusberg, S. J. (2019). The role of metaphor in communication and thought. *Language and Linguistics Compass*, *13*(5), 1-18. https://doi.org/10.1111/lnc3.12327

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Van Herwegen, J., Dimitriou, D., & Rundblad, G. (2013). Development of novel metaphor and metonymy comprehension in typically developing children and Williams syndrome. *Research in Developmental Disabilities*, *34*(4), 1300–1311. https://doi.org/10.1016/j.ridd.2013.01.017

Wechsler, D. (2011). Wechsler Abbreviated Scale of Intelligence—Second Edition. *APA PsycTests*. https://doi.org/10.1037/t15171-000

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span.

Acta Psychologica, *115*(2–3), 167–183.

https://doi.org/10.1016/j.actpsy.2003.12.005