frontiers | Frontiers in Communication

TYPE Review
PUBLISHED 07 February 2025
DOI 10.3389/fcomm.2025.1448463

OPEN ACCESS

EDITED BY
Antonio Benítez-Burraco,
University of Seville, Spain

REVIEWED BY
Rob Davies,
Lancaster University, United Kingdom
Hannah Lutzenberger,
Stockholm University, Sweden

*CORRESPONDENCE
Derya Çokal
✉ dcokal@uni-koeln.de

RECEIVED 13 June 2024
ACCEPTED 20 January 2025
PUBLISHED 07 February 2025

CITATION
Çokal D, Bishop J, Torregrossa J, Patterson C, Grice M, Wehrle S, Lialiou M, Repp S, Seeliger H, Eisenbeiß S, von Heusinger K, Vogeley K and Schumacher PB (2025) Individual differences in discourse management.
Front. Commun. 10:1448463.
doi: 10.3389/fcomm.2025.1448463

Individual differences in discourse management

Derya Çokal[1]*, Jason Bishop[2], Jacopo Torregrossa[3], Clare Patterson[1], Martine Grice[1], Simon Wehrle[1], Maria Lialiou[1], Sophie Repp[1], Heiko Seeliger[1], Sonja Eisenbeiß[1], Klaus von Heusinger[1], Kai Vogeley[1] and Petra B. Schumacher[1]

[1]University of Cologne, Cologne, Germany, [2]The City University of New York, New York, NY, United States, [3]Goethe University Frankfurt, Frankfurt, Germany

Every utterance in discourse we produce arises from the interaction of numerous cognitive functions, such as semantic memory, where we store the meanings of words, executive function and working memory as required for maintenance of a discourse goal, and social cognitive abilities, such as mind-reading capacity as required for tuning what we say to what others know or believe. In this way, a single utterance potentially integrates very different cognitive capacities into a basic discourse processing unit. This suggests that discourse processing and management is a very rich phenomenon that requires a multidimensional approach. We propose that a model of discourse management is comprised of three primary components that interact synergistically: (i) dynamicity, (ii) predictability, and (iii) meta-representationality. Cognitive functions play a pivotal role in the underlying processes, contributing to the development and unfolding of discourse. Understanding the correspondence between individual differences in discourse management (i.e., discourse perception and production) and cognitive functions can shed light on the intricate relationship between language and cognition in discourse management, as well as the appropriate psychometric measures to address this complex interaction. This narrative review presents aspects of discourse management, psychometric measures to comprehensively address these aspects. We close with a discussion of challenges and open questions.

KEYWORDS

discourse, cognitive functions, autism, bilingualism, individual differences, prosody, anaphora processing, social cognition

1 Introduction

Effective discourse management across both spoken and written modalities involves the interplay of three cognitive key components: dynamicity, predictability, and meta-representationality (Newen and Vogeley, 2003; Vogeley and Bartels, 2011; von Heusinger and Schumacher, 2019). These components are fundamental for navigating both spoken and written discourse, including engaging in conversations with multiple participants, resolving ambiguity, and adapting to continuously unfolding information in discourse representation. In this narrative review, we argue that these components are central to discourse management (e.g., the construction and updating of a discourse), and are intimately linked to basic cognitive functions such as working memory (WM) and executive functions (EFs), as well as social cognition capacities such as mind-reading capacity (e.g., also referred to as social reasoning, mentalizing or theory of mind). Ultimately, this link between discourse management processes and underlying cognitive functions may account for individual differences in the choice and interpretation of expressions and constructions in discourse. Understanding this connection matters

because it allows us to identify measurable cognitive processes underlying discourse management. By doing so, we can determine how individual differences in cognitive functions such as WM, EFs, and social cognition shape interactions in discourse. This insight not only deepens our understanding of language use but also informs the design of psycho/neurolinguistic experiments and new psychometric tools to assess and predict communication effectiveness in diverse contexts (cf. Brysbaert's (2024) tutorial on designing and evaluating tasks to measure individual differences). Such insights are foundational for advancing linguistic theory and they also hold practical relevance for education, clinical diagnosis, innovative communication technologies, and speech therapy (e.g., for people with communication problems). Establishing this link enables the identification of cognitive profiles that influence language use, helping us better understand how individuals process, adapt, and engage in discourse, particularly in diverse social and cognitive contexts.

Individual differences refer to variations among individuals in discourse management, particularly in relation to key cognitive aspects such as WM, EFs, social cognition capacities, personality traits and states, and/or personal interests. Specifically, in our research within the framework of *Prominence in Discourse*, we have encountered significant individual variability at the discourse level. Furthermore, a very recent article by Dietrich et al. (2024) underscores the importance of investigating individual differences in future research. Specifically, they suggest examining how discourse processing strategies correlate with variations in working memory (e.g., working memory load affects intact and violated presupposition). A similar call is echoed in the recent book *Individual Differences in Anaphora Resolution* by Fotiadou and Tsimpli (2023), which underscores the scarcity of research on individual differences at the discourse level and the cognitive processing involved. Accordingly, we provide a narrative review of components of discourse management, present a model of discourse management that relies on prominence relations, and examine how these components may be influenced by individual traits. This is the first attempt to connect a discourse management model with cognitive functions, demonstrating how individual differences can inform, and how research on individual differences can be informed by such a model. Our focus encompasses both spoken (conversation) and written (text-based) discourse management.

To lay out the basis for our review, in the following sections we first provide a brief overview of previous review articles on individual differences in linguistics. Next, we outline our methodology. Third, we discuss the three fundamental components of discourse management that participants utilize. We then identify the cognitive processes and resources involved in discourse management. Following that, we review psychometric measures of individual differences in cognitive functions and social cognition capacities. To establish the idea that discourse management can be assessed by psychometric tests addressing cognitive functions, we then review case studies that specifically investigate individual differences in discourse management, covering prosody, neurotypicality, online language comprehension, and bilingualism. Finally, we discuss the challenges involved in testing and analyzing individual differences data, and we highlight directions for future studies.

1.1 Previous reviews on individual differences

To our knowledge, there are no reviews specifically addressing individual variation in discourse management that encompasses both spoken and written modalities. Therefore, to situate our review, we first survey previous review articles on individual differences across sub-disciplines of linguistics. In a review article, Kidd et al. (2018) examine recent research in psycholinguistics, and propose that linguistic and psycholinguistic theories have downplayed the existence of meaningful variation in language use among individuals. Specifically, they discuss how a focus on individual differences offers critical insights into language acquisition and processing. Specifically, they emphasize the importance of language experience in language acquisition, processing, attainment, as well as the underlying structure of the language system. Therefore, they show the relationship and interaction between language and cognitive systems and thus put forth three key imperatives for theories and future studies: (1) The *existence imperative*, which calls for theories of acquisition and processing of language to predict meaningful individual differences in language; (2) The *environmental imperative* that highlights the role of input, emphasizing the relationships between language abilities and the amount and nature of linguistic input individuals receive; and (3) The *architectural imperative*, which urges theoreticians to explain the connections between linguistic subsystems and underlying cognitive processes.

Building upon this research, Matthews et al. (2018) review studies on *pragmatic skills* in children, and highlight their associations with pragmatic abilities, formal language (vocabulary and grammar), social cognition [Theory of Mind (ToM)], and EFs. In addition, Matthews et al. (2018) point out that very few studies have examined such relationships because there are relatively few tests of individual differences in pragmatic skills that have good psychometrics in terms of reliability, validity, and distributional properties. They also demonstrate that the cognitive processes underlying pragmatic inferences are often not well-defined, making it difficult to determine why specific pragmatic tasks should correlate with broader cognitive measures. While their approach examines components like formal language, social cognition, and EFs to understand discrete abilities such as irony comprehension and contingent conversation, our approach emphasizes discourse management. We focus on the interplay of dynamicity, predictability, and meta-representationality as essential for managing discourse representation, resolving ambiguity, and processing unfolding information. Moreover, while the authors point to methodological limitations (i.e., sample size, good quality measures for sufficient variance in pragmatics, clear information on processing, and controls for theoretical important covaraites like formal language and non-verbal IQ) and propose developmental taxonomies of pragmatic skills, we provide a framework for discourse management across spoken and written modalities and propose how components of discourse management interact with cognitive and social abilities.

Expanding the focus to a phonological perspective, Yu and Zellou (2019) observe that individual variation is evident in most phonological behaviors. However, few studies address the heterogeneity of language processing and production, as the focus

tends to remain on group-level patterns. They provide an overview of individual variability within phonology, examining how linguistic patterns can differ among individuals. They compare a traditional population-level approach, which relies on experimental manipulations and aggregated responses, with an individual-differences approach that explores naturally occurring variability. Their review suggests that language variation and sound change may be influenced by individual differences in how phonological information is stored and processed. Interindividual variation has also been examined with respect to the use of various phonetic parameters demonstrating that individual listeners and speakers weigh certain acoustic cues differently (e.g., Bishop, 2016; Baumann and Winter, 2018; Bishop et al., 2022; Lorenzen and Baumann, 2024).

While previous reviews have explored individual differences within specific language domains—such as pragmatic abilities among children, and phonological processing—our review offers an integrated view, since several aspects of discourse management across spoken and written modalities are considered at the same time. Responding to Kidd et al.'s (2018) call, we provide a model of discourse management comprised of three key components and propose that the link between these discourse management components and cognitive functions may explain individual variability at the discourse level. Our review builds on existing literature by specifically examining the relationship between discourse management and cognitive functions across different contexts and populations. In particular, we cover discourse related prosody in production and comprehension, online discourse comprehension, and reference management among bilinguals, thereby bringing a new dimension to discussions of individual differences at the discourse level. This novelty arises both from the domains we explore and the approach we adopt.

1.2 Methods

Given that our review is a narrative review with critical reflection, our methodology involves the targeted selection of studies that provide both theoretical and empirical insights that illustrate the fundamental components of discourse management and their interaction with cognitive functions. Psychometric measures were chosen based on their frequency of use within sub-domains of linguistics and their relevance to exploring the relationship between cognitive functions and discourse management. We chose case studies where we observe individual variation at the discourse level, where evidence exists for the utility of components as distinct constructs, or where the predictive validity of cognitive measures related to discourse management was present. Specifically, we included studies from our own labs and other published research highlighting individual differences in discourse management.

1.3 Components of discourse management and their interaction with cognitive functions

To show the significance of the discourse components and their contributions to discourse management and cognition, we consider a written text that describes a scenario involving two individuals, a cello player and a critic, at a concert.

> (1)
> (a) The cello player wants to impress the critic. (b) However, he is asleep at the wheel. (c) The cello player is completely frustrated about the situation.

In (1a), the cello player and his feelings about the critic are introduced. When the personal pronoun *he* is read or heard by readers and listeners in (1b), the pronoun is immediately associated with the cello player since the interpretation is guided by features of the antecedent expression (i.e., cello player in (1a) represents the first-mentioned, subject, more agentive entity) and is thus the prominent antecedent for the personal pronoun *he* (cf. e.g., the first-mentioned entity advantage in pronoun resolution in Gernsbacher and Hargreaves, 1988; Carreiras et al., 1995). However, as the discourse unfolds (i.e., "asleep at the wheel"), it becomes evident that *he* refers to the critic, not the cello player because this is the more plausible scenario at this stage of the discourse. Hence referential discontinuity (i.e., a shift from *the cello player* to *the critic*), necessitates a dynamic update of the discourse representation, yielding co-reference of the personal pronoun in (1b) with the critic. As (1) shows, dynamicity is crucial for constructing and managing a coherent discourse representation (e.g., Kamp, 1981; Lascarides and Asher, 2007; Dekker, 2012; von Heusinger and Schumacher, 2019). This process requires the active processing of incoming information, integrating it with existing knowledge, and tracking information structures within discourse. In addition, it involves constructing and updating mental models and discourse structure (cf. mental models in Garnham, 2001; Johnson-Laird, 1983; Johnson-Laird and Garnham, 1980; Kintsch and van Dijk, 1978).

As seen above, predictability also plays a major role in understanding the next discourse units (Friston, 2010; Bornkessel-Schlesewsky and Schumacher, 2016; von Heusinger and Schumacher, 2019). Based on the information in (1a), readers and listeners predict that the narrative will be about the cello player's actions and/or the critic's response. However, in (1b), the idiomatic expression 'asleep at the wheel' suggests negligence by the critic, prompting readers and listeners to predict the potential consequences of this negligence. In (1c), readers and listeners conclude the cello player's disappointment is a consequence of the critic's behavior. In such cases, they anticipate the narrative's development, facilitating the creation of a mental representation of discourse. Numerous theoretical frameworks have modelled this predictive behavior in discourse with respect to accessibility, activation, or attentional centering (see von Heusinger and Schumacher, 2019 for an overview). We adopt the prominence framework, where varies cues contribute to discourse management. Traditionally, grammatical and thematic cues have been the focus of research on prominence-lending cues. However, our proposal also extends to multimodal cues, such as (pointing) gesture, head nods or other non-manual cues.

In addition to these two components of dynamicity and predictability, meta-representationality come into play as readers and listeners consider how language mirrors thoughts and emotions (Alcalá-López et al., 2018, 2019 for the perspective of neuroscience; Lin et al., 2018; Hinterwimmer, 2019). For efficient discourse management, they need to understand verbal and non-verbal

cues—including linguistic expressions, tone of voice, facial expressions, and body posture. Difficulties in meta-representationality can cause discourse management problems, such as misunderstanding the interlocutor's intended meaning and purpose. This, in turn, can result in communication failures, including miscommunication, lack of coherence in replies, and misuse of expressions. For instance, the sentence 'the cello player wants to impress the critic' conveys the cello player's desire for approval. In this case, meta-representation also prompts the readers and listeners to activate the meaning of the figurative expression 'asleep at the wheel'. In (1c), meta-representations, by means of which we model the inner experience or the mental phenomena of another person, enable the readers and listeners to understand the emotional and cognitive states of the cello player (Baron-Cohen et al., 1985; Frith and Frith, 1999; Frith, 2012; Gallagher et al., 2002; Newen and Vogeley, 2003 for psychopathology). They recognize that the cello player's disappointment is linked to the critic's absent-mindedness (i.e., 'asleep at the wheel').

These components – dynamicity, predictability, meta-representationality – are not entirely independent or orthogonal; rather, they function together to facilitate discourse management. During the interaction of these three components, several cognitive processes are assumed to underlie these functions on a cognitive level (see 2.2 & 3 below). Working Memory (WM) and Executive Functions (EFs) are involved in both dynamicity and predictability (see the predictive capacity of WM in Baddeley, 2022; Rönnberg et al., 2022a, 2022b). WM covers various memory types, including short-and long-term, explicit, implicit, episodic, and visuospatial memory and contributes to active retention (Baddeley, 2000, 2012). EFs[1] are essential abilities to control one's attention updating and suppress irrelevant information to keep track of entities (executive attention in Engle, 2018; attentional control mechanism in Miyake et al., 2000). Specifically, the roles of WM and EFs in prediction and postdiction (i.e., including inference-making and repair) are supported by the case studies from cognitive hearing science (Rönnberg et al., 2021). Mismatches between predictions and linguistic inputs trigger postdictions, enabling further access to WM (e.g., long term memory) (Rönnberg et al., 2021).

Social cognition is an empirical indicator of meta-representationality. With this in mind, it is based on perspective taking, theory of mind (ToM), empathy, language proficiency, and language experience. The use and development of social cognition is typically explicit and reliant on WM and EFs (Carlson et al., 2002; Truscott and Smith, 2022), as well as language experience. Representations of social cognition, which involve the coactivation of associated representations or "schemas," are closely related to WM (Truscott and Smith, 2022). Moreover, the combination of inhibition and WM is central to the relation between EFs and ToM. This is supported by studies on developmental trajectories and interrelations between EFs and ToM (Carlson et al., 2002, 2004), reviews of the relation between EFs and ToM (Pineda-Alhucema et al., 2018), and research on the relation between EFs, ToM, and autism (Hemmers et al., 2022).

Language experience also facilitates the interpretation of linguistic and subtle social cues, such as implicatures or prosody, which are essential for inferring emotions and intentions. Some individuals perform well with respect to these cognitive functions including social cognition capacities, while others may struggle, leading to individual differences in discourse management. Several psychometric tests (e.g., *complex span task* for working memory; the *Hick task* for processing speed) are used to measure these cognitive functions and social cognition capacities (We discuss these in section 2.). Therefore, we propose that discourse management can be assessed by psychometric tests addressing cognitive functions. In the next section we provide a summary of commonly used psychometric tests for assessing individual differences related to the components of the discourse management proposed above.

2 Cognitive factors in discourse management

As was discussed in the introduction, the following cognitive functions and social cognition abilities are essential to ensure coherent and contextually relevant discourse management:

(a) Working memory (WM), which allows for the storage and manipulation of information within the mental workspace while working on other levels of the discourse model (i.e., essential for dynamicity, predictability, & meta-representationality).
(b) Executive functions (EFs), which facilitate the inhibition, updating, and shifting of information (i.e., critical for dynamicity, predictability, & meta-representationality).
(c) Social cognition, which involves the mental representation of others' minds and the social aspects of discourse and which is closely related to language experience, including language and reading proficiency (i.e., critical for dynamicity, predictability, & meta- representationality).

Variations in any of these cognitive functions can lead to individual differences in discourse management and processing across spoken and written modalities. In the following section, we present frequently used psychometric tests in sub-domains of linguistics, designed to measure the critical components of our discourse model.

2.1 Dynamicity and predictability

As seen in Example (1) above, effective discourse management necessitates dynamicity, which requires the active processing of incoming information, integrating it with existing knowledge, and tracking linguistic cues within discourse. This dynamic process is facilitated by WM, which holds and updates information, and EFs, which maintains attention on unfolding information. Predictability plays a crucial role in this process, allowing individuals to predict upcoming information or events based on the current context and prior knowledge. Continuously generating predictions about what might come next guides attention, suppresses irrelevant information,

1 EF also plays a crucial role in connecting or facilitating communication between the frontal lobes and the long-term memory systems-specifically those involved in language comprehension and production (Coolidge and Wynn, 2022).

facilitates comprehension, and fills in gaps in the discourse [see review on measuring individual differences in WM and attentional control and their contribution to language comprehension in Burgoyne et al. (2022)]. Following Frischkorn et al. (2022), we view EFs as cognitive processes related to attention during goal-oriented processing, subdivided into three subsequent categories: inhibition, shifting, and updating. By making predictions, recalling/holding previously mentioned information in WM, and updating these predictions through EFs, individuals make inferences and adjust their discourse representations accordingly. Therefore, WM and EFs are essential for handling the dynamic nature of discourse management and making accurate predictions during communication. In addition to WM and EFs, processing speed significantly influences individual differences in dynamic discourse, representation and prediction (Huettig, 2015; Huettig and Janse, 2016). Faster processing allows for quicker integration of new information and more accurate predictions, while slower processing may hinder the ability to update and generate accurate predictions. Below, we provide an overview of frequently used psychometric measures.

As mentioned earlier, working memory capacity (WMC) is associated with various cognitive skills, including attention, concentration, and cognitive flexibility. The most frequently used instrument for estimating WMC is the family of so-called *complex span tasks* (Daneman and Carpenter, 1980; Turner and Engle, 1989; Engle et al., 1992; Oberauer et al., 2000; Miyake et al., 2001; Wilhelm et al., 2013). A widely used alternative complex span task, which requires storage and recall of stimulus items and includes an additional processing task, is the *reading span task* introduced by Daneman and Carpenter (1980). The task is sometimes criticized for its similarity to reading comprehension (see Daneman and Hannon for further discussions). Subsequent research adapted the original reading span task and developed different versions (LaPointe and Engle, 1990; Whitney et al., 2001; van den Noort et al., 2008).

Other WMC paradigms that entail both storage and processing include an *n-back task* and a *running memory span task* (Kirchner, 1958). In the n-back task, participants are presented with a sequence of stimuli for which they must indicate whether the current stimulus in the sequence matches a stimulus that was presented *n* times back. Similarly, in a running memory span task, participants are presented with a list consisting of an unpredictable number of items. As soon as the list ends, participants must recall items from the end of the list. Both paradigms require an ability to maintain and continuously update information in WM. In a recent fMRI study, the *n-back task* has been used to examine the general feature of brain processing for predictions of upcoming events (e.g., see studies that examine WMC for predictions with individuals with hearing loss in Rönnberg et al., 2013, 2022a, 2022b; Sörqvist et al., 2016). In addition to WMC, measuring executive processes appears to be an important tool in understanding individual behavior in discourse management. For this purpose, commonly used tasks in experimental psychology are the *Stroop task* (Stroop, 1935), the *Flanker task* (Eriksen and Eriksen, 1974), and the *Simon task* (Simon et al., 1981). All these standard tasks assess how conflicting information influences both response choice and response speed. While the Stroop task is assumed to be cross-linguistically valid (Matthews et al., 2018), the Stroop and Flanker measures, in particular, do not appear to be correlated (Hedge et al., 2018). In addition, the Stroop and Flanker inhibition tasks might be unrelated due to low correlations in larger datasets and thus contradicting the concept of a unified account (Rouder and Haaf, 2019) (see potential analyses in 4.1 below).

Processing speed significantly influences individual differences in discourse management. As discussed in Frischkorn et al. (2022), the term "processing speed" is quite diverse as it can refer to mental speed, information processing speed as well as attention speed. However, there is consensus that processing speed is essential to better understand individual differences. Commonly used tasks that attempt to measure processing speed include the *Hick*, the *Sternberg*, and the *Odd-Man-Out* paradigms (Hick, 1952; Sternberg, 1969; Frearson and Eysenck, 1986), and the Letter and Pattern Comparison tasks (Salthouse, 1991; Salthouse and Babcock, 1991).

2.2 Meta-representationality

The meta-representationality in discourse management can be assessed through measures of social cognition, language proficiency, and reading proficiency. Below, we provide an overview of psychometric measures to evaluate these capacities.

2.2.1 Social cognition

We should note that there are more psychometric measures on social cognition (Msika et al., 2024) but we, here, refer to the ones broadly used in linguistic studies. In research on pragmatic abilities, individual measures of social cognition have garnered increasing attention since they may reflect Theory of Mind, or perspective taking, capacity and social reasoning impacting discourse management and the processing of speaker's meaning. For instance, the Autism Spectrum Quotient Questionnaire (AQ) (Baron-Cohen et al., 2001b) is a self-assessment questionnaire measuring autistic traits in adults. It consists of five subscales that assess communication skills, social skills, imagination, attention to local detail, and capacity to switch attention. The communication subscale taps into pragmatic abilities directly. Some examples of the items from this subscale are "Other people frequently tell me that what I have said is impolite, even though I think it is polite," "I find it hard to 'read between the lines' when someone is talking to me," and "I am often the last to understand the point of a joke." The score on the overall test – or sometimes just a subscale score – has been used to account for individual variation (Bishop, 2012). Although originally developed as screening tool for the detection of autism spectrum disorder, the AQ has been used in the recent past to measure communicative abilities with the understanding that autism can serve as a model disturbance for social cognition. The AQ communication subscale and measures are related to the *Reading the Mind in the Eyes Test* (Baron-Cohen et al., 2001a). Two personality-related measurements assess empathy (e.g., *Interpersonal Reactivity Index* [IRI] by Davis, 1980, *Empathy Quotient* by Baron-Cohen and Wheelwright, 2004). For example, the IRI includes four sub-scales—perspective taking, fantasy, emphatic concern, and personal distress—on which participants rate themselves.

2.2.2 Language proficiency

Language proficiency tests evaluate an individual's language experience to process linguistic cues in discourse. When including different groups of speakers/listeners in studies on language processing (adults and children with different linguistic profiles, e.g., speaking the language as their first or a second language, literate or illiterate, etc.),

there will likely be considerable variation, attributable to different levels of proficiency in the language (Tremblay, 2011; Park et al., 2022). As we will discuss in Section 3.1.3, proficiency may also be connected to the use of one language across different contexts over time. Crucially, language proficiency may affect a speaker's ability to develop meta-representations of related to the form and function of linguistic units used in discourse (Karmiloff-Smith, 1992). Such representations are essential for effectively managing the components of discourse management. Where proficiency-based individual variation is itself under investigation, researchers should use standardized tests and consider using several different proficiency measures.

It is outside the scope of this article to review the many different ways of assessing proficiency, as well as advantages and disadvantages of these methods (for reviews of proficiency reporting practices, see Tremblay, 2011; Hulstijn, 2012; Park et al., 2022). Here we briefly present three commonly used proficiency measures suitable for psycholinguistic research because of their availability, ease of administration, and validity. It should be noted that these tests may be more reflective of reading proficiency than oral proficiency; researchers should consider other proficiency measures such as *Elicited Imitation* if oral proficiency is relevant (Torregrossa et al., 2024). *LexTALE* represents a test designed to measure vocabulary knowledge and general proficiency in advanced leaners (Lemhöfer and Broersma, 2012). Participants are presented with words and non-words one by one and they must make an untimed lexical decision (e.g., *Is this a word in language x or not?*). While the test was originally designed and validated for English (Lemhöfer and Broersma, 2012), Dutch, German and French versions are also available (see lextale.com for German and Dutch, and Brysbaert, 2013).The advantages of LexTALE are that it is quick to administer, freely available, validated (English version), and pre-programmed for several experimental software programs. While LexTALE appears to tap lexical knowledge and speed of lexical retrieval, test scores map well onto more standardized general proficiency test results. In addition, LexTALE scores have been shown to correlate with knowledge of discourse phenomena (Wetzel et al., 2020).

In a *Cloze test*, participants are given a text that contains one-word gaps and the task is to correctly fill in the gaps. Carefully designed Cloze tests are thought to reflect morphosyntactic, lexical and discourse competencies (Tremblay, 2011). Cloze tests have many of the same advantages as LexTALE: quick and easy to administer, and freely available in several languages including Arabic, Bengali, English, German, Dutch, Japanese, Korean, Turkish, French, Portuguese, and Spanish (Norris, 2018). The disadvantage of a Cloze test is text complexity. In addition, researchers should ensure tests are appropriate for the precise population they wish to test and validate against standardized proficiency tests (Tremblay and Garrison, 2010; Norris, 2018).

Though criticized for its subjectivity and lack of independent verification (Hulstijn, 2012; Park et al., 2022), self-reports of proficiency are widely used in psycholinguistic studies where proficiency is not part of the main research question. Participants are usually asked to rate their second language (L2) proficiency in four areas: reading, writing, speaking, and listening. Ratings are usually on a 5, 7, or 10-point scale. There is evidence that participants are good at estimating their own proficiency and that self-reported proficiency correlates well with standardized proficiency tests (Ross, 1998; Marian et al., 2007; Lemhöfer and Broersma, 2012).

2.2.3 Reading proficiency

Reading proficiency tests are used to evaluate an individual's ability to understand and interpret texts. While these tests primarily focus on reading comprehension skills, they can also provide insights into a person's capacity to consider an interlocutor's perspective in discourse. Effective reading comprehension often requires readers to infer authors' intentions, understand characters' perspectives, and grasp underlying themes.

A number of studies converge on a set of individual differences or latent factors that influence reading comprehension success. For instance, Ahmed et al. (2016) demonstrated that knowledge and vocabulary contribute indirectly to reading comprehension through inferencing. Follmer and Sperling (2018) showed the role of EFs in processing of referential cohesion. While shifting predicted reading comprehension for lower cohesion, updating was essential for higher cohesion. Notably, metacognitive monitoring and reading strategies contributed to reading comprehension. Kulesz et al. (2016) highlighted vocabulary and background knowledge as dominant predictors of comprehension, with genre emerging as the most influential text feature, whereas other skills like word reading, reading fluency, and working memory, were less significant in accounting for reading comprehension. These studies highlight the critical roles of vocabulary, background knowledge, and inferential reasoning in reading comprehension, while emphaszing the need to consider text demands and reader-specific skills (e.g., see Perfetti and Stafura (2013) for an over-arching theoretical syntheses of findings on individual differences in reading).

While reading comprehension is influenced by these higher-level cognitive processes, reading proficiency also encompasses low-level skills, such as word decoding (e.g., Hamilton et al., 2016) and visual word recognition (Yap et al., 2011). It entails integrating new information with existing knowledge, predicting upcoming information or events based on context, and constructing and updating discourse representations (see Tighe and Schatschneider, 2016 for a review). Some decoding tasks include (1) a Phonological Decision Task (Bell and Perfetti, 1994); (2) Non-word Naming Task (Bell and Perfetti, 1994); and (3) Orthographic Decision Task (e.g., DEAL vs. DEEL) (Bell and Perfetti, 1994). While reading proficiency may affect reading speed and word identification, it also reflects upon a more general (medium-independent) capacity relating to discourse management, enabling individuals to navigate complex communicative interactions successfully.

In psycholinguistics, reading proficiency is further assessed through standardized tests that probe vocabulary knowledge and reading comprehension, such as the Nelson Denny Reading Test (Brown, 1960; Nelson, 1991), which consists of a multi-choice vocabulary and a reading comprehension test. In addition, print exposure measures such as the Author Recognition Test (ART) (Stanovich and West, 1989) are used to investigate individual differences in reading experience (Acheson et al., 2008; Arnold et al., 2018; Johnson and Arnold, 2021; Langlois and Arnold, 2020; Scholman et al., 2020; Wetzel et al., 2020; Zufferey and Gygax, 2020). In the ART, participants are shown a list of real and fake (i.e., foils) author names and are asked to identify the author names they recognize. Participants earn one point for every real author they tick, and one point is deducted for every foil ticked. The inclusion of foils provides additional validity to the test because participants cannot simply obtain a high score by ticking every author. The

original version contained 50 authors and 50 foils. The ART serves as a proxy for print exposure, providing an indirect measure of how much participants read. While print exposure is correlated with reading proficiency, it is more specifically a measure of reading experience rather than a direct measure of proficiency. Print exposure has been shown to correlate with lower level (word recognition) and higher level (vocabulary knowledge, reading comprehension) reading skills (Mol and Bus, 2011; Moore and Gordon, 2015).

Obviously, the author list must be current and relevant to the target participant group, which is why the test is often adapted to the area of study. This has resulted in a proliferation of test versions (e.g., Acheson et al., 2008; Martin-Chang and Gould, 2008; Moore and Gordon, 2015), including versions targeting particular age- and language-groups (e.g., Cunningham and Stanovich, 1990; Grolig et al., 2017, 2020; Schroeder et al., 2016; Zufferey and Gygax, 2020). In addition, author names are often taken from recent bestseller lists for a particular region. Compared to more standardized cognitive tests, this tailoring of the ART to a participant group may make comparison of performance across studies challenging.

3 Case studies in discourse management

The studies in this section provide insights into why it is essential to assess the components of discourse management (i.e., dynamicity, predictability, and meta-representationality) with psychometric measures addressing cognitive functions (i.e., WM, EFs) and social cognition capacities.

3.1 Individual differences in prosody for discourse management

We review some recent findings from prosodic prominence to illustrate individual differences across different domains and modalities, emphasizing their role in dynamicity, predictability, and meta-representationality in discourse management. To begin with, speakers choose different prosodic means to signal different speech acts (Repp, 2020; Repp and Seeliger, 2020; Seeliger and Repp, 2023). For instance, prosodic prominence marking can dynamically signal specific elements in discourse, such as the verb participle in (2a) or the subject d-pronoun "der" in (2b), to mark the speech acts question and exclamation, respectively.

> (2)
>
> (a) Question speech act—accent on the verb participle: Hat der geSCHRIen? (Has he screamed?)
>
> (b) Exclamations—accent on der/he: Hat DER geschrien! ((Boy), did he scream!) (Repp and Seeliger, 2020, p. 1).

These findings highlight the dynamicity of prosodic prominence as speakers adaptively make choices in discourse. In addition, production studies reveal that the placement of the so-called *exclamative accent*, which is a speech-act-specific prominent accent whose placement is fairly independent of information structure, like the accent on the subject *d*-pronoun "der" in (2b), is individual-specific. While some speakers consistently accent the pronoun "der," while others consistently accent the finite verb "hat" (Repp and Seeliger, 2020). Such individual-specific patterns highlight the dynamic nature of prosodic prominence marking in discourse. Thus, interlocutors need to predict such prosodic prominence markers and update their representation to ensure coherent discourse.

Having examined production variability in exclamatives, we now turn to rejecting questions, which further underscore the interplay of dynamicity, predictability, and meta-representationality. Similarly, inter-individual differences in prosodic prominence marking are also evident in rejecting questions in German. Rejecting questions are questions that speakers use to express their disbelief in something the contextual evidence suggests (English: "Surely, you do not want to steal this paper"). Production studies show that speakers differ in their preferences for rising vs. falling contours for these questions (Repp and Seeliger, 2020). This variability demonstrates how speakers dynamically adapt prosodic prominence marking to signal pragmatic meaning. At the intra-individual-level, speakers exhibit considerable variability, which does not seem to depend on particular discourse contexts (Repp and Seeliger, 2020). Since such variability does not align with discourse context, the role of meta-representationality and dynamicity in prosodic prominence production becomes essential for managing discourse successfully. Further research is needed to understand the cognitive and meta-representation reasons behind such variations, as well as how these variations impact discourse management.

Compared with production, less is known about how individual differences affect the perception of prosodic prominence, though emerging evidence strongly implicates socio-cognitive abilities. Much of what we know so far has come from work exploring the role of social cognition capacities in various perception or comprehension-related tasks involving prosodic manipulations. For example, Bishop (2012, 2017) found that listeners with higher scores on the AQ (a socio-cognitive measure previously mentioned above), show weaker sensitivity to prosodic prominence patterns in tasks requiring predictability of prosodic cues, such as a cross-modal lexical decision task. Importantly, sensitivity is not related to overall AQ scores but to subscale scores targeting communication (the communication subscale), pointing to a nuanced link between socio-cognitive abilities and prosodic processing. This suggests that listeners' meta-representational abilities—such as the capacity to integrate prosodic cues into higher-order social reasoning—might play a central role in prosodic perception. In subsequent work, the AQ communication subscale and measures closely related to it—such as the *Reading the Mind in the Eyes Test* (Baron-Cohen et al., 2001a). The pragmatic language subscale of the *Broad Autism Phenotype Questionnaire* (Hurley et al., 2007)—have also been shown to predict sensitivity to prominence patterns in tasks in which listeners must explicitly rate or identify prominence in auditory stimuli (Bishop, 2016; Bishop et al., 2020) or use it to recover information structure for sentence completion (Hurley and Bishop, 2016). Notably, listeners exhibit differential sensitivity depending on social cognition measures, pointing to the role of social cognition abilities in discourse management.

Moreover, in two studies designed to prime prosodic structure (analogous to syntactic priming), Jun and Bishop (2015a, 2015b) argue that individuals with more autistic-like or weaker communication skills exhibit less robust implicit representations of prosodic prominence (i.e., subvocal) during silent reading. Such

social-cognitive deficits results in weaker meta-representationality of discourse management, which impacts individuals' ability to predict, adapt, and manage discourse effectively.

These findings underscore the importance of considering individual differences in socio-cognitive abilities when examining prosodic prominence in discourse management. Therefore, several empirical findings indicate that scores on measures like the AQ communication subscale can predict some of the variation listeners exhibit in response to prosodic prominence patterns in spoken, and perhaps even internally-generated utterances. At present, however, it remains unclear just what aspect of social cognition these measures reflect. We assume that this is related to a deficit in implicit social cognition that is used during ongoing conversations, whereas in offline studies without personal participation (e.g., written language), persons with ASD can effectively participate in and manage activities (Zimmermann et al., 2021).

Interestingly, and possibly consistent with this interpretation, Jouravlev et al. (2020) present fMRI data that show AQ communication scores (but not scores on the other subscales of the AQ) are associated with less hemispheric lateralization during neurotypical language processing in the direction of increased right hemisphere activity. As we discuss further below, our understanding of this sort of variation will benefit from the employment of more rigorous psychometric methodologies. Tentative evidence pointing in this direction includes recent studies showing that measures of empathy—a construct broadly related to perspective taking but arguably distinct from ToM—also seem to predict variation in sensitivity to prosody and intonation (Esteve-Gibert et al., 2020; Orrico and D'Imperio, 2020; Arvaniti et al., 2022). Above all, the findings suggest that a different, or perhaps more general socio-cognitive construct, might in fact be the relevant one. Whatever the correct circumscription of this relationship between social cognition and sensitivity to prosody turns out to be, understanding the implications for inter-individual variation in discourse representation and management will be an important task.

Finally, in conversational interactions, melodic speech patterns and backchannels further reveal the role of socio-cognitive abilities in discourse management. For example, in a corpus with Autism Spectrum Disorder (ASD) speakers, most speakers produced a more melodic intonation style than non-autistic speakers (Wehrle et al., 2022). In another production study on backchannels (e.g., "mmhm, okay"), many ASD speakers produced rising intonation contours regardless of the backchannel type, indicating a less flexible mapping of intonation contours to backchannel type (Wehrle et al., 2024b). The ASD group also produced fewer of the expected level contours on filled pauses (e.g., "uhm"), although once again this pattern did not hold true for all speakers (Wehrle et al., 2024a). These studies point to the importance of inter-individual variability in clinical groups such as speakers with autism (Bishop, 2012, 2017; Grice et al., 2016; Grice et al., 2023; Wehrle, 2023). These findings suggest that individuals with different socio-cognitive profiles may exhibit varying sensitivities to melodic speech patterns and backchannels, impacting their ability to predict and interpret discourse dynamicity effectively.

In summary, studies on individual differences in prosodic prominence marking, melodic speech patterns, and backchannels highlight how components of discourse management interact with cognitive and social abilities.

3.2 Individual differences in discourse comprehension

Referentiality plays a critical role in discourse management, bridging non-verbal cognitive and linguistic skills to facilitate communication. Anaphoric expressions, such as pronouns in sentence comprehension, engage processes related to dynamicity, predictability, and meta-representationality.

Anaphoric expressions require continual updating of mental representations as new information becomes available, enabling readers to anticipate upcoming text based on syntactic constructions. For instance, resumptive pronouns (RPs) are pronouns that appear in a position where a gap would normally be predicted, as the italicized RP *it* in (3a).[2] Constructions that block a dependency between a gap and an extracted entity are referred to as islands.

> (3)
>
> (a) RP in the non-island condition: Jane liked the magazine that the hairdresser had talked about *it* before going to the salon.
>
> (b) RP in the island condition: Jane liked the magazine that the hairdresser [RC who had talked about it before going to the salon] bought (Çokal and Sturt, 2022).

The reader promptly establishes a gap in (3a) right after the embedded verb (such as "talked about"), creating a dependency with the so-called filler "that." However, when the explicit direct object (such as *it*) is encountered in the input, it conflicts with initial prediction, resulting in a disruption in processing, which leads to a dynamic updating of discourse representation to resolve the conflict. Moreover, dependency formation is sensitive to islands and thus a filled gap effect is not observed in (3b), where the RP *it* appears inside a strong island (i.e., relative clause). In two eye-tracking reading experiments, Çokal and Sturt (2017) demonstrate that filler-gap dependency formation interacts with individual reading skills, which were measured using the *Nelson-Denny Reading Test* (Brown, 1960). The interaction pattern shows that high-skilled readers' total reading times increase in the non-island condition in (3b) compared to low-skilled readers. Low-skilled readers' processing times for non-island condition in (3a) and island condition in (3b) do not differ. The fact that low-skilled readers' reading times do not differ across conditions suggests that maintaining multiple interpretations is less robust in individuals with lower reading proficiency. While Çokal and Sturt's (2017) study shows a relationship between reading proficiency and filler-gap dependency formation in islandhood conditions, the intricate relationship between processing of RPs and cognitive functions (particularly WM & EFs) has not yet been investigated. Further research is needed to understand whether the observed processing disadvantage for pronouns in island configurations implies that individuals with higher WM and EFs may be able to predict syntactic complexity and maintain referential dependencies across clause boundaries.

2 In fact, they are ungrammatical in English, but some studies have shown that they can facilitate sentence comprehension (Beltrama and Xiang, 2016) and thus their acceptability (Hofmeister and Norcliffe, 2013).

While Çokal and Sturt highlight the role of individual reading skills in processing syntactic dependencies, other studies explore how print exposure influences the predictability and dynamicity of referential prominence at the discourse level. Arnold et al. (2018) investigate ambiguous pronoun processing and its relationship to print exposure (measured by the ART). This study found a higher ART score correlated with a higher subject-bias for ambiguous pronoun comprehension in English but not WM and ToM. Importantly, Arnold et al. (2018) propose that increased print exposure strengthens participants' prediction about which referents (e.g., subjects versus objects) are likely to be re-mentioned. This ability reflects a dynamic process of updating mental representations, as individuals must maintain a strong candidate referent for ambiguous pronouns while simultaneously suppressing irrelevant ones. The study by Arnold et al. (2018) suggests the interplay between dynamicity, predictability and EFs, showing how print exposure enhances readers' ability to anticipate referents and suppress irrelevant candidates. However, to our best knowledge, Arnold et al. have not tested the role of executive functions (EFs) in the processing of ambiguous pronouns. Examining the interaction between EFs and print exposure could shed further light on how individuals dynamically manage and predict referential candidates in larger discourse contexts. This remains an important avenue for future research.

The role of cognitive functions is further supported by electrophysiological evidence, which underscores individual differences in processing ambiguous pronouns. Individual differences with respect to ambiguous pronouns (e.g., "Jennifer Lopez told Madonna that she had too much money.") have been tested using ERPs and the Reading Span Task (Nieuwland and Van Berkum, 2006). While ambiguous pronouns elicit a sustained, frontal negative shift (Nref) relative to non-ambiguous pronouns at the group-level, the size of this effect positively correlates with the reading span and contextual bias. Individuals who are sensitive to ambiguous pronouns have higher reading skills than low skilled readers. This seems to indicate that high-span readers demonstrate greater dynamicity in anticipating and maintaining ambiguous referents, while low-span readers adopt the most plausible interpretation immediately, reflecting weaker dynamicity and predictability. These findings provide neurocognitive evidence for the interaction between WM and components of discourse management (i.e., dynamicity and predictability) in processing referential expressions.

To understand how cognitive functions and social cognition abilities influence referential resolution, Vogelzang et al. (2021) present a computational model. Vogelzang et al. (2021) modeled individual differences using the Cognitive Architecture-Adaptive Control of Thought-Rational Model (ACT-R Model) (Anderson and Anderson, 2007). Their findings specifically demonstrate that effective discourse management involves understanding the perspectives and intentions of others, as well as incorporating dynamicity and predictability. In Italian – the language considered by the authors—null pronouns are used to refer to a prominent entity in the previous discourse, signalling topic continuation, whereas overt pronouns are used to refer to a less prominent entity, indicating a shift of topic from a prominent to a less prominent entity. Individual differences in processing of pronouns have been observed in both children and adults. With this in mind, Vogelzang et al. (2021) simulate the processing of pronouns by adults. Vogelzang et al. (2021) add high WM load to their model as an additional component through a spreading mechanism. Thus, the selection of discourse topic is predicted to be influenced by WMC. Based on this framework, Vogelzang et al. (2021) model adults' processing of null/overt pronouns, specifying the effects of WMC and processing speed. According to the default mechanisms of the ACT-R Model, a referent's base-level activation is determined by how recently and frequently it has appeared in the discourse. This activation gradually diminishes over time but is boosted when the referent reappears. Vogelzang et al. (2021) formulate hypotheses about children's processing of pronouns and then create a simulation for children. According to their model, null pronoun interpretation is influenced by WMC, which appears to depend on the dynamicity of discourse. On the other hand, for adult processing, overt pronoun interpretation is influenced by processing speed, which is related to perspective taking. However, children are not "adult-like" in their interpretation of null pronouns due to a lack of perspective taking. Their results demonstrate that discourse processing is influenced by WMC, while social cognition capacities are influenced by processing speed.

In conclusion, the processing of anaphoric expressions (e.g., resumptive pronouns, personal pronouns, null or overt pronouns) is the interaction between components of discourse management and cognitive and social abilities. For example, print exposure strengthens predictability (Arnold et al., 2018), reading proficiency or WM enhances dynamic updating of representations (Çokal and Sturt, 2017; Nieuwland and Van Berkum, 2006), and processing speed supports meta-representational strategies like perspective-taking (Vogelzang et al., 2021). Further investigation is warranted to elucidate the specific contributions of these cognitive functions to the management of larger discourse units. Specifically, knowledge of the weight of each cognitive factor in pronoun resolution in a larger discourse context is needed.

3.3 Interspeaker and intraspeaker variation in bilingual reference management

Bilingual speakers provide a unique perspective for understanding how components of discourse management interact with cognitive and social abilities, owing to their variability in language experience and cognitive and social-cognitive profiles [e.g., working memory (WM), executive functions (EFs), and theory of mind (ToM)]. Specifically, the analysis of bilingual reference production allows us to examine inter- and intraspeaker variations in discourse management. Let us consider interspeaker variation first. Bilingual speakers may differ with respect to meta-representationality—specifically language experience and ToM variables. In general, bilinguals may be dominant in one language, depending on the contexts of language use. Furthermore, bilingual speakers may vary from each other in their cognitive profiles (i.e., WM, EFs, mentalizing). The same variation in cognitive profiles can be observed among monolingual speakers. However, bilingual speakers represent a privileged viewpoint for understanding how social cognition capacities and cognitive profiles interact with each other in discourse processing, as also observed among monolingual speakers. For instance, Torregrossa et al. (2021) investigate how discourse updating skills, language proficiency, and their interaction predict the use of overspecified and underspecified referring expressions in bilingual children. To achieve this, they analyzed the narrative production of 125 bilingual Greek children (aged

7;2–13;1) who spoke Greek as their heritage language and Albanian, English, or German as their societal languages. The study measured the frequency of overspecified (full nouns) and underspecified (null-subjects/clitics) referring expressions, which differ in ambiguity. The children also completed vocabulary tasks in both languages and a 2-back task to assess updating skills (a component of EFs). Parents provided information about language use through a questionnaire, from which a dominance index was calculated based on vocabulary scores and language exposure. The study examined how updating skills, dominance, and their interaction predicted the use of referring expressions in Greek, with separate analyses for children dominant in Greek or the societal language. In Torregrossa et al. (2021), the use of overspecified forms was predicted by their discourse updating skills: the lower their skills, the more overspecified forms they produce. This highlights how cognitive flexibility and updating contribute to referential predictability by allowing children to balance specificity and ambiguity effectively. For children dominant in the societal language, language proficiency emerges as the primary predictor of overspecified forms. Interestingly, language combination also impacts referential strategies: Greek-German children produce more overspecified forms compared to Greek-Albanian children. This variation underscores the role of meta-representationality in integrating linguistic and social experience to manage discourse. For underspecified forms, the findings reveal an inverse pattern. The more balanced the two languages, the more underspecified forms are produced. This suggests that balanced bilingualism enhances flexibility in choosing referring expressions, reflecting the dynamicity of bilingual discourse management.

A contribution by Torregrossa et al. (2023) relates to the understanding of intraspeaker variation in bilingual children's metalinguistic knowledge concerning the use of referring expressions in discourse, considering the effect of language mode (Torregrossa et al., 2023). The study provided the sentences in three modes: (1) Italian sentences in isolation (one-language mode, one sentence); (2) Italian sentences paired with their correct counterparts (one-language mode, two sentences), and (3) Italian sentences paired with their Greek counterparts (bilingual mode). The results show that children perform better in the bilingual mode (Italian alongside Greek), particularly in identifying errors in referring expressions. This finding highlights the dynamicity in bilingual discourse management, where the interaction between two languages enhances meta-representational awareness. The ability to identify errors more accurately in bilingual contexts suggests that bilingual children can draw on cross-linguistic cues to strengthen referential predictability and maintain coherence in discourse management. This study underscores the role of language context in shaping meta-representationality related to language use and discourse management.

By examining inter- and intraspeaker variation, these bilingual studies reveal how linguistic factors, cognitive function, and social cognition abilities including language experience shape discourse management strategies among bilinguals. The findings emphasize the following:

- Dynamicity: Bilinguals choose referring expressions based on cognitive flexibility, discourse updating skills, and cross-linguistic interaction.
- Predictability: Linguistic dominance and proficiency strengthen the ability to predict upcoming discourse, choose correct referential expressions, and detect referential errors that mismatch the previous context.
- Meta-representationality: Meta-linguistic awareness and cross-linguistic comparison enhance the ability to manage errors and integrate referential strategies dynamically across contexts.

Importantly, these results suggest that discourse management can be assessed through psychometric tools targeting cognitive function and social-cognitive abilities.

4 Discussion

It is theoretically meaningful to understand the interplay between individual differences in discourse management (i.e., dynamicity, predictability, and meta-representationality), cognitive functions (e.g., WM, EFs), and social cognition capacities, as well as to examine how individuals' characteristics interact with discourse management. However, this research topic has not yet received enough attention. Nevertheless, a few studies have utilized psychometric measures to explore this aspect (see section 3.). From these studies we have learned: (1) Social cognition abilities and prosody: Two social cognition measures, namely the sub-scales of the Autism Quotient (AQ) and measures of empathy, can predict individual differences in subjects' sensitivity to prosodic prominence patterns in discourse management (Bishop, 2012, 2017); (2) Variability in discourse management among clinical groups (e.g., ASD speakers): socio-cognitive abilities effect discourse management by influencing individuals' sensitivity to melodic speech patterns and backchannels. (3) Cognitive functions, reading proficiency and referential ambiguity resolution: Individuals with greater WM, faster processing speed, and higher reading skills often excel in tasks that require them to consider multiple potential antecedents and rapidly resolve pronoun ambiguity in a coherent and efficient manner. (4) Social cognition abilities in children: Children were not "adult-like" in their interpretation of null pronouns due to a lack of perspective taking. In addition, processing of null and overt pronouns is influenced by WM, while perspective taking is influenced by processing speed and ToM. (5) Bilingual meta-representationality: The role of metalinguistic knowledge in reference production in bilinguals' discourse, and its relationship with social cognition measures such as ToM, vocabulary knowledge, and language proficiency has been shown in Torregrossa et al. (2021, 2023). (6) Dynamicity and predictability in discourse management: Children who are skilled at updating their mental representation of a sentence as it unfolds (a skill associated with WM and cognitive flexibility) seem to be able to adapt their language production and discourse management strategies, avoiding over- or underspecification (Torregrossa et al., 2021).

In light of these findings, we claim that discourse management can be assessed by psychometric tests addressing cognitive functions. The interplay between discourse components and cognitive functions occurs as language and visual input enter the cognition and language system. Language and visual input activate WM and EFs, which correspond to each other. The information from these cognitive

functions combined with language and visual input in discourse is used to generate predictions about upcoming input. If there is a rapid match between language/visual input and predictions, then discourse management is dynamically updated. However, when there is a mismatch between language/visual input and predictions, postdiction (including inference-making and repair) commences, which feeds back into new predictions, creating a dynamic interacting system (Rönnberg et al., 2013, 2022a, 2022b). There is also an interaction between dynamicity and predictability with meta-representationality. Consequently, postdiction may or may not be successfully accomplished depending on individuals' meta-representationality. With psychometric measures addressing cognitive functions, we can identify individuals' discourse abilities and provide insights into discourse management.

4.1 Which psychometric measures of cognitive functions can be used to assess discourse management?

Our review shows that several psychometric measures have been used in previous studies on discourse management: WM (e.g., complex span tasks, which include some kind of additional mental task in-between exposure and recall), processing speed, Nelson Denny Reading Test, Author Recognition Test, Language Proficiency Test, and two social-cognition measures (namely communication-related sub-scales of measures like the Autism Quotient (AQ) and the Broad Autism Phenotype Questionnaire, and measures of empathy).

Interestingly, previous studies have not thoroughly explored the role of EFs (i.e., inhibition). In fact, inhibition – filtering/suppressing irrelevant thoughts while updating mental representations, maintaining coherence, making predictions, logical progression in discourse, and inhibiting one's own perspective while focusing on contextual information – is a fundamental cognitive and social skill that helps individuals navigate the complexities of discourse management.

We propose that conscientiousness (i.e., one of the basic personality/processing types in some models of personality types) might assess individual differences as the propensity to be self-controlled, responsible to others, and rule abiding (Roberts et al., 2009, 2012). The level of perfectionism that individuals might possess can be linked to their conscientiousness in discourse management. To our knowledge, such a connection between perfectionism and discourse management has not yet been investigated. The perfectionism model of Gaudreau and Thompson, 2010; Gaudreau (2013), which has two interacting sub-traits of perfectionism: (a) personal standards perfectionism (PSP) and (b) evaluative concerns perfectionism (ECP) can be used to address this gap (see Drizinsky et al., 2016 for perfectionism-related variations in behavioral and neurophysiological correlates of error recognition). Including such measurement in future studies will serve as a control to demonstrate that the correlation between cognitive functions – EFs and WM – is not solely attributable to the general cognitive functions assessed but is also related to personality traits. In other words, some people might consider a "good enough" completion of a task to be sufficient, whereas others might aim for perfection. Thus, attitudes can also contribute to individual variation.

4.2 What challenges arise when testing and analyzing individual differences data?

There are some challenges to consider before conducting a study on individual differences in discourse management (cf. Boogert et al., 2018 for major challenges), during data analysis (Tremblay, 2011), and when developing and evaluating experiments (Brysbaert, 2024). Below, we discuss these challenges (which include inter-correlation measures and multicollinearity, and sampling concerns) and provide potential analysis approaches to address them. We emphasize that addressing these challenges needs to be done in future studies.

Many cognitive constructs (e.g., WM, ART, reasoning) are intercorrelated, complicating the attribution of outcomes to specific predictors (Dubois et al., 2018; James et al., 2018). High correlations across tasks can lead to inconsistent findings (e.g., Martin et al., 2020; Freed et al., 2017). A major challenge in this context is multicollinearity, where highly correlated predictors (e.g., WM and ART) complicate regression analysis (Cohen et al., 2003). Increasing sample size reduces random noise but does not fully resolve multicollinearity because correlations between variables often persist. Alternative approaches include reducing predictor correlations, combining redundant predictors, or selecting variables based on theoretical relevance. Statistical adjustments like ridge regression can help but are insufficient on their own. Thoughtful variable selection and careful experimental design are critical to ensuring model clarity and minimizing multicollinearity (Cohen et al., 2003).

The next challenge is sample size and amount of variation in the data. If there is a higher level of measurement uncertainty (i.e., more variability or error in measurements), then more samples will be required to achieve reliable results. For instance, if we want to test for a correlation between a prominence in discourse management (such as the "rate of pronouns and full NPs") and a variable representing a cognitive construct (such as the "EFs"), there needs to be enough variation in both variables. However, if the range of variation across individuals in one or both variables is small—either because individuals do not actually vary much along the dimension in question or because the sample size is too small—this would not allow for a robust test of individual differences. This limitation might also result in a type II error (i.e., a false negative). In addition, traditional approaches have attempted to classify individuals as "do-ers" (who show an effect) and "do not-ers" (who do not show an effect) based on individuals' cognitive processing style and effect (Haaf and Rouder, 2019; Hedge, 2021). However, Haaf and Rouder (2019) argue that this classification approach is inadequate. Oversight of such variability can mask meaningful findings or anomalies (e.g., spike-and-slab model below).

Sampling representativeness is often incongruous with research goals in individual differences. For instance, studies on individual differences do not focus on obtaining a representative sample of participants across different socio-cognitive profiles. Instead, recruited participants are often relatively homogeneous—such as easily accessible participants, monolingual/bilingual speakers of a specific language, or university students raised in similar linguistic environments, and exposed to controlled or consistent language input.

Since no single measure can provide a pure and error-free measure of a construct such as WM, researchers are advised to follow several strategies. To capture variation between individuals in the use of components of discourse, it is essential to employ multiple experiment

tasks with a strong theoretical perspective across phonology, discourse processing, and speech production. Then, multiple trials (i.e., test–retest) should be conducted to examine whether the tasks capture stable characteristics (see Boogert et al., 2018 for further methodological issues). It should be noted that there could be additional sources of error in a test–retest reliability context due to fluctuations in mood, health, and training effects (Hedge, 2021). Alternatively, instead of multiple trials, researchers can use multiple measures for each construct to "… assess the degree of common variance between them and use composite scores within a construct" (James et al., 2018, p. 5). For instance, a composite score can be created by administering the Stroop task (Stroop, 1935), the Flanker task (Eriksen and Eriksen, 1974), and/or the Simon task (Simon et al., 1981) (cf. James et al., 2018 for syntactic processing). However, Rouder and Haaf (2019) demonstrated that Stroop and Flanker inhibition tasks seem unrelated and contradict the unified concept of inhibition (see Friedman and Miyake, 2017 for the details of unified account). These findings highlight the need to critically assess whether a single task (or multiple tasks for that matter) can accurately capture individual differences in discourse management.

To address these challenges, advanced statistical methods like Structural Equation Modelling (SEM) and Psychometric Network Modelling are particularly effective. SEM handles inter-correlations among cognitive functions (e.g., WM, fluid intelligence, ART) while accounting for measurement error and multidimensional relationships (Schreiber et al., 2006; Marsh et al., 2014). On the other hand, Psychometric Network Modelling (Goring et al., 2021; Kovacs and Conway, 2016) conceptualizes cognitive abilities as networks of interconnected processes rather than latent variables. Observed variables are nodes, with partial correlations modeled as edges. This isolates direct relationships while filtering out shared variance, providing a modern alternative to latent variable approaches.

To avoid oversimplified classifications (e.g., "do-ers" vs. "do not-ers"), Vogelzang and Torregrossa et al. (2024) also propose cluster analysis to identify behavioral profiles, which are then compared on cognitive and linguistic dimensions. The spike-and-slab model (Haaf and Rouder, 2019) further refines this by identifying mixtures of effects (e.g., "some do, some do not") and determining whether tasks capture a single dimension or distinct processes. To deal with effect sizes within tasks, correlations of these effects across tasks, and noise from finite trials (i.e., trial noise), Rouder and Haaf (2019) propose hierarchical models (see *Bayes Factor*). This is necessary when aggregating scores across trials for individuals introduces trial-level noise, which contaminates measures like effect size, reliability, and correlation, resulting in artificially low values. This method provides cleaner estimates of individual performance and enables robust evaluation of correlations between tasks. All these statistical methods highlight the need for careful experimental design (e.g., many trials per condition) to accurately measure and interpret individual differences.

As seen above, the way data is modelled (e.g., using a specific statistical model) can impact the interpretation of relationships between measures (see conflict in modelling in Freed et al., 2017; Goring et al., 2021). It should be kept in mind that different modelling approaches can lead to different theoretical conclusions about cognitive functions (Goring et al., 2021). However, computational models generally require participant numbers in the hundreds rather than the dozens. Unlike mixed-effects regression, these methods help researchers avoid suffering from multicollinearity among independent variables, allowing for a better understanding of the associations between cognitive functions and discourse management. To address some of these challenges, researchers have begun conducting multi-lab studies to tackle difficulties related to replication and methodological issues (e.g., manybabies: https://manybabies.org/).

4.3 How critical are the challenges to the validity of the case studies reported above?

The challenges outlined in section 3 will be considered here in relation to their potential impact on the validity of the case studies reported above, as well as future research directions. Construct validity in the case studies is ensured through the use of psychometric measures targeting well-defined components of discourse, such as WM, EF, and social cognition capacities. It should be noted that while the challenges highlight areas for improvement, they do not invalidate the claims themselves. Instead, they underscore the need for continued methodological refinement and replication of results.

As discussed in section 4.2, while inter-correlation and multicollinearity between psychometric measures often pose challenges by leading to redundant predictors and reducing interpretability, these issues may be less critical for many of the case studies discussed here. Most of these case studies utilize only a single psychometric measure to link a specific cognitive construct to discourse management. For instance, the Nelson Denny Test in Çokal and Sturt (2017) assesses reading skills, the Reading Span Task in Nieuwland and Van Berkum (2006) measures WMC, while socio-cognitive measures – such as the AQ (Bishop, 2012; Bishop, 2017) and the Broad Autism Phenotype Questionnaire (Bishop, 2016; Bishop et al., 2020)—examine social cognition capacities. The use of a single psychometric measure in these studies minimizes the risk of overlapping variance between multiple predictors within the same analysis, thereby simplifying interpretability and reducing concerns about inter-correlation and multicollinearity.

However, multicollinearity can arise if two or more independent variables in a model share a significant amount of variance, potentially obscuring their unique contributions. This challenge is particularly relevant in studies employing multiple psychometric measures, such as Arnold et al. (2018) and Vogelzang et al. (2021). In Arnold et al. (2018), the predictors (e.g., WMC, ToM, & ART) did not correlate. Therefore, intercorrelation might not be an issue. It is important to note that this review article is not a meta-analysis but rather a discussion of individual case studies to highlight that discourse management can be assessed by psychometric tests addressing cognitive functions.

In turn, it should be noted that sample size and amount of variation in the data could critically influence the validity of the case studies discussed above. For instance, prosody studies we reported above involve sample sizes ranging between 84 and 160 participants (e.g., Bishop, 2012; Bishop, 2016; Bishop et al., 2020; Hurley and Bishop, 2016; Jun and Bishop, 2015a, 2015b). Lab-based experiments, such as those utilizing EEG or eye-tracking, often include smaller sample sizes, ranging from 31 to 40 participants (e.g., Nieuwland and Van Berkum, 2006; Çokal and Sturt, 2017), with exceptions such as Arnold et al. (2018) involving 72 and 60 participants, and Torregrossa et al. (2021), which examined 125 bilingual Greek children. Similarly, clinical studies frequently rely on small participant groups due to the specificity of their inclusion criteria (Grice et al., 2016; Wehrle, 2023; Wehrle et al., 2024a). These sample size constraints are often related to the requirements of such studies, which involve recruiting participants who meet precise criteria and

compensating them for their time. While small sample sizes and constrained variation can reduce studies' generalizability, such challenges may be less critical for studies with robust designs. For example, two-by-two designs, as employed by Nieuwland and Van Berkum (2006) and Çokal and Sturt (2017), allow for more controlled comparisons, mitigating issues of variability in the data. This suggests that while sample size and variability may impact generalizability, the methodological rigor of specific studies can somewhat offset these challenges.

The representativeness of the samples depends on the specific purpose of each study. The reported case studies predominantly include university students, which limits the generalizability of findings to populations with different socio-cognitive profiles. Similarly, clinical studies often focus on a specific group (e.g., high-functioning individuals with ASD), whose recruitment and accessibility present challenges compared to more commonly studied groups, such as native speakers in phonological experiments. Furthermore, studies involving bilingual participants tend to target specific language pairs aligned with the research goals. While representativeness is less critical for the current focus of this review (i.e., proposing the groundwork for the intricacies of discourse management and its relationship to cognitive functions influencing individual variation rather than population-level effects), it is relevant for future research.

The statistical approaches employed in these case studies, including mixed-effects models with binary categorization (e.g., "doers vs. non-doers" or "high vs. low skilled"), can also be refined. Advanced methods, such as cluster analysis, structural equation modeling (SEM), or psychometric network modeling, are well suited to disentangle relationships between predictors and outcomes while accounting for trial-level variability and measurement error. Such methods are particularly valuable in studies examining complex relationships among constructs like WM and EFs. Moving beyond binary classifications and employing robust statistical models that integrate multiple psychometric tasks for targeted cognitive functions would enhance the validity and interpretability of findings in the case studies.

Overall, while methodological challenges such as inter-correlation, multicollinearity, sample size, representativeness, and statistical approaches are areas for refinement, the case studies reviewed here maintain validity within their specific contexts and offer valuable insights into the link between cognitive functions and discourse management, paving the way for more comprehensive and generalizable future research.

4.4 Future directions

Our review poses several open questions that further studies need to address:

(1) Do we observe the same individual differences at the single-sentence level for larger discourse-level contexts/units?
(2) What is the relation between executive functions (EFs) and processing ambiguous pronouns in discourse management within a larger context?
(3) How does the perception, processing, and interpretation of prosodic prominence vary across individuals? Do we observe inter-individual variation in perception as we do in production?
(4) Can discourse management success be predicted by psychometric measures associated with cognitive functions? Can such success be a biomarker for clinical groups?
(5) To what extent can the proposed framework be extended to visual–spatial modalities such as sign language and gesture? What additional considerations are necessary for such an extension?

In conclusion, in this review we have laid the groundwork for a more in-depth exploration of this nuanced and underexplored field, opening the door to further investigations into the intricacies of discourse management and its relationship to cognitive functions influencing individual variation. While our focus has been on spoken and written modalities, we recognize the importance of multimodality—particularly the visual–spatial domain—and emphasize the need for future research to explore how our proposed framework applies to sign language and gesture. The proposed measures and considerations outlined here offer valuable insights and pathways for future research, promising to enhance our understanding of this complex and evolving subject area.

Author contributions

DÇ: Conceptualization, Investigation, Project administration, Writing – original draft, Writing – review & editing. JB: Writing – original draft, Writing – review & editing. JT: Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing. CP: Writing – original draft, Writing – review & editing. MG: Writing – original draft, Writing – review & editing. SW: Writing – original draft, Writing – review & editing. ML: Writing – original draft, Writing – review & editing. SR: Writing – original draft, Writing – review & editing. HS: Writing – original draft, Writing – review & editing. SE: Writing – original draft, Writing – review & editing. KH: Funding acquisition, Writing – original draft, Writing – review & editing. KV: Funding acquisition, Writing – original draft, Writing – review & editing. PS: Conceptualization, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) -Project-ID 281511265 -SFB 1252 Prominence in Language -at the University of Cologne, Department of German Language and Literature I, Linguistics.

Acknowledgments

We would like to express our gratitude to Tuncay Tosun for his assistance in gathering and organizing the references.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Acheson, D. J., Wells, J. B., and MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behav. Res. Methods* 40, 278–289. doi: 10.3758/BRM.40.1.278

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., and Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemp. Educ. Psychol.* 45, 68–82. doi: 10.1016/j.cedpsych.2016.02.002

Alcalá-López, D., Smallwood, J., Jefferies, E., van Overwalle, F., Vogeley, K., Mars, R. B., et al. (2018). Computing the social brain connectome across systems and states. *Cereb. Cortex* 28, 2207–2232. doi: 10.1093/cercor/bhx121

Alcalá-López, D., Vogeley, K., Binkofski, F., and Bzdok, D. (2019). Building blocks of social cognition: Mirror, mentalize, share? *Cortex* 118, 4–18. doi: 10.1016/j.cortex.2018.05.006

Anderson, J. R., and Anderson, J. R. (2007). How can the human mind occur in the physical universe? Oxford, NY: Oxford University Press.

Arnold, J. E., Strangmann, I. M., Hwang, H., Zerkle, S., and Nappa, R. (2018). Linguistic experience affects pronoun interpretation. *J. Mem. Lang.* 102, 41–54. doi: 10.1016/j.jml.2018.05.002

Arvaniti, A., Gryllia, S., Zhang, C., and Marcoux, K. (2022). 'Disentangling emphasis from pragmatic contrastivity in the English H* ~ L+H* contrast', in. *Proc. SpeechProsody* 2022, 837–841. doi: 10.21437/SpeechProsody.2022-170

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/s1364-6613(00)01538-2

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Baddeley, A. (2022). "Working memory and the challenge of language" in *The Cambridge Handbook of Working Memory and Language*. (Eds.) Schwieter J. W, Wen Z., (Edward), Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46. doi: 10.1016/0010-0277(85)90022-8

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175. doi: 10.1023/b:jadd.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001a). The "Reading the mind in the eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001b). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/a:1005653411471

Baumann, S., and Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *J. Phonet.* 70, 20–38. doi: 10.1016/j.wocn.2018.05.004

Bell, L. C., and Perfetti, C. A. (1994). Reading skill: some adult comparisons. *J. Educ. Psychol.* 86, 244–255. doi: 10.1037/0022-0663.86.2.244

Beltrama, A., and Xiang, M. (2016). Unacceptable but comprehensible: the facilitation effect of resumptive pronouns. *Glossa* 1:29. doi: 10.5334/gjgl.24

Bishop, J. (2012). WPP, no.111: focus, prosody, and individual differences in "autistic" traits: evidence from cross-modal semantic priming. *UCLA Work. Papers Phonetics* 111, 1–26.

Bishop, J. (2016) 'Individual differences in top-down and bottom-up prominence perception', in. Proceedings of Speech Prosody, pp. 668–672.

Bishop, J. (2017). Focus projection and prenuclear accents: evidence from lexical processing. *Lang. Cogn. Neurosci.* 32, 236–253. doi: 10.1080/23273798.2016.1246745

Bishop, J., Kuo, G., and Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: evidence from rapid prosody transcription. *J. Phon.* 82:100977. doi: 10.1016/j.wocn.2020.100977

Bishop, J., Zhou, C., Antolovic, K., Grebe, L., Hwang, K. H., Imaezue, G., et al. (2022). Brief report: autistic traits predict spectral correlates of vowel intelligibility for female speakers. *J. Autism Dev. Disord.* 52, 2344–2349. doi: 10.1007/s10803-021-05087-5

Boogert, N. J., Madden, J. R., Morand-Ferron, J., and Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170280. :. doi: 10.1098/rstb.2017.0280

Bornkessel-Schlesewsky, I., and Schumacher, P. (2016). Towards a neurobiology of information structure. Oxford: Oxford University Press.

Brown, J. I. (1960). The Nelson-Denny Reading test. Oxford: Houghton Mifflin (The Nelson-Denny Reading Test).

Brysbaert, M. (2013). Lextale_FR a fast, Free, and Efficient Test to Measure Language Proficiency in French. *Psychol. Belgica* 53:23. doi: 10.5334/pb-53-1-23

Brysbaert, M. (2024). Designing and evaluating tasks to measure individual differences in experimental psychology: a tutorial. *Cogn. Res.* 9:11. doi: 10.1186/s41235-024-00540-2

Burgoyne, A. P., Martin, J. D., Mashburn, C. A., Tsukahara, J. S., Draheim, C., and Engle, R. W. (2022). "Measuring individual differences in working memory capacity and attention control and their contribution to language comprehension" in The Cambridge handbook of working memory and language. eds. Schwieter J. W, Wen Z. (Edward), Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

Carlson, S. M., Mandell, D. J., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40, 1105–1122. doi: 10.1037/0012-1649.40.6.1105

Carlson, S. M., Moses, L. J., and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant Child Dev.* 11, 73–92. doi: 10.1002/icd.298

Carreiras, M., Gernsbacher, M. A., and Villa, V. (1995). The advantage of first mention in Spanish. *Psychon. Bull. Rev.* 2, 124–129. doi: 10.3758/BF03214418

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences. *3rd* Edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Çokal, D., and Sturt, P. (2017). 'Resumptive pronouns and active dependency formation', in *CUNY sentence processing conference*. Boston: OSF.

Çokal, D., and Sturt, P. (2022). The real-time status of strong and weak islands. *PLoS One* 17:e0263879. doi: 10.1371/journal.pone.0263879

Coolidge, F. L., and Wynn, T. (2022). "The evolution of working memory and language"in *The Cambridge Handbook of Working Memory and Language*. (Eds.) Schwieter J. W, Wen Z., (Edward), Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

Cromley, J. G., Snyder-Hogan, L. E., and Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: a domain-specific test of the direct and inferential mediation model of reading comprehension. *J. Educ. Psychol.* 102, 687–700. doi: 10.1037/a0019452

Cunningham, A. E., and Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: a quick measure of reading experience. *J. Educ. Psychol.* 82, 733–740. doi: 10.1037/0022-0663.82.4.733

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog Select. Doc. Psychol* 10:85.

Dekker, P. J. E. (2012). Dynamic semantics. Dordrecht, Heidelberg: Springer.

Dietrich, S., Seibold, V. C., and Rolke, B. (2024). Discourse comprehension and referential processing: effects of contextual distance and semantic plausibility on presupposition processing. *Lang. Cogn.* 16, 2032–2054. doi: 10.1017/langcog.2024.45

Drizinsky, J., Zülch, J., Gibbons, H., and Stahl, J. (2016). How personal standards perfectionism and evaluative concerns perfectionism affect the error positivity and post-error behavior with varying stimulus visibility. *Cogn. Affect. Behav. Neurosci.* 16, 876–887. doi: 10.3758/s13415-016-0438-z

Dubois, J., Galdi, P., Paul, L. K., and Adolphs, R. (2018). 'A distributed brain network predicts general intelligence from resting-state human neuroimaging data', *philosophical transactions of the Royal Society of London. Series B Biol. Sci* 373:20170284. doi: 10.1098/rstb.2017.0284

Engle, R. W. (2018). Working memory and executive attention: a revisit. *Perspect. Psychol. Sci.* 13, 190–193. doi: 10.1177/1745691617720478

Engle, R. W., Cantor, J., and Carullo, J. J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 972–992. doi: 10.1037/0278-7393.18.5.972

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Esteve-Gibert, N., Schafer, A. J., Hemforth, B., Portes, C., Pozniak, C., and D'Imperio, M. (2020). Empathy influences how listeners interpret intonation and meaning when words are ambiguous. *Mem. Cogn.* 48, 566–580. doi: 10.3758/s13421-019-00990-w

Follmer, J. D., and Sperling, R. A. (2018). Interactions between reader and text: contributions of cognitive processes, strategy use, and text cohesion to comprehension of expository science text. *Learn. Individ. Differ.* 67, 177–187. doi: 10.1016/j.lindif.2018.08.005

Fotiadou, G., and Tsimpli, I. M. (Eds.) (2023). Individual Differences in Anaphora Resolution: An introduction. *Individual Differences in Anaphora Resolution: Language and cognitive effects*, Benjamins Publishing Company, 1–20.

Frearson, W., and Eysenck, H. J. (1986). Intelligence, reaction time (RT) and a new "odd-man-out" RT paradigm. *Personal. Individ. Differ.* 7, 807–817. doi: 10.1016/0191-8869(86)90079-6

Freed, E. M., Hamilton, S. T., and Long, D. L. (2017). Comprehension in proficient readers: the nature of individual variation. *J. Mem. Lang.* 97, 135–153. doi: 10.1016/j.jml.2017.07.008

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

Frischkorn, G. T., Wilhelm, O., and Oberauer, K. (2022). Process-oriented intelligence research: a review from the cognitive perspective. *Intelligence* 94:101681. doi: 10.1016/j.intell.2022.101681

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 2213–2223. doi: 10.1098/rstb.2012.0123

Frith, C. D., and Frith, U. (1999). Interacting minds: a biological basis. *Science* 286, 1692–1695. doi: 10.1126/science.286.5445.1692

Gallagher, H. L., Jack, A. I., Roepstorff, A., and Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814–821. doi: 10.1006/nimg.2002.1117

Garnham, A. (2001). Mental models and the interpretation of anaphora. Hove: Psychology Press.

Gaudreau, P. (2013). The 2 × 2 model of perfectionism: commenting the critical comments and suggestions of Stoeber (2012). *Personal. Individ. Differ.* 55, 351–355. :. doi: 10.1016/j.paid.2013.03.021

Gaudreau, P., and Thompson, A. (2010). Testing a 2 × 2 model of dispositional perfectionism. *Personal. Individ. Differ.* 48, 532–537. doi: 10.1016/j.paid.2009.11.031

Gernsbacher, M. A., and Hargreaves, D. J. (1988). Accessing sentence participants: the advantage of first mention. *J. Mem. Lang.* 27, 699–717. doi: 10.1016/0749-596X(88)90016-2

Goring, S. A., Schmank, C. J., Kane, M. J., and Conway, A. R. A. (2021). Psychometric models of individual differences in reading comprehension: a reanalysis of Freed, Hamilton, and Long (2017). *J. Mem. Lang.* 119:104221. doi: 10.1016/j.jml.2021.104221

Grice, M., Krüger, M., and Vogeley, K. (2016). Adults with Asperger syndrome are less sensitive to intonation than control persons when listening to speech. *Cult. Brain* 4, 38–50. doi: 10.1007/s40167-016-0035-6

Grice, M., Wehrle, S., Krüger, M., Spaniol, M., Cangemi, F., and Vogeley, K. (2023). Linguistic prosody in autism spectrum disorder—an overview. *Lang. Linguist. Compass* 17:e12498. doi: 10.1111/lnc3.12498

Grolig, L., Cohrdes, C., and Schroeder, S. (2017). Der Titelrekognitionstest für das Vorschulalter (TRT-VS): Erfassung des Lesevolumens von präkonventionellen Leserinnen und Lesern und Zusammenhänge mit Vorläuferfertigkeiten des Lesens. [the title recognition test for kindergarteners (TRT-VS): assessment of preconventional readers' print exposure and its relations to precursors of reading.]. *Diagnostica* 63, 309–319. doi: 10.1026/0012-1924/a000186

Grolig, L., Tiffin-Richards, S. P., and Schroeder, S. (2020). Print exposure across the reading life span. *Read. Writ. Interdiscip. J.* 33, 1423–1441. doi: 10.1007/s11145-019-10014-3

Haaf, J. M., and Rouder, J. N. (2019). Some do and some don't? Accounting for variability of individual difference structures. *Psychon. Bull. Rev.* 26, 772–789. doi: 10.3758/s13423-018-1522-x

Hamilton, S., Freed, E., and Long, D. L. (2016). Word-decoding skill interacts with working memory capacity to influence inference generation during Reading. *Read. Res. Q.* 51, 391–402. doi: 10.1002/rrq.148

Hedge, C. (2021). Qualitative individual differences are useful, but reliability should be assessed and not assumed. *J. Cogn.* 4:48. doi: 10.5334/joc.169

Hedge, C., Powell, G., and Sumner, P. (2018). The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* 50, 1166–1186. doi: 10.3758/s13428-017-0935-1

Hemmers, J., Baethge, C., Vogeley, K., and Falter-Wagner, C. M. (2022). Are executive dysfunctions relevant for the autism-specific cognitive profile? *Front. Psych.* 13:886588. doi: 10.3389/fpsyt.2022.886588

Hick, W. E. (1952). On the rate of gain of information. *Q. J. Exp. Psychol.* 4, 11–26. doi: 10.1080/17470215208416600

Hinterwimmer, S. (2019). Prominent protagonists. *J. Pragmat.* 154, 79–91. doi: 10.1016/j.pragma.2017.12.003

Hofmeister, P., and Norcliffe, E. (2013). The core and the periphery: Data-driven perspectives on syntax inspired by Ivan a. Stanford, CA: CSLI publications.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Res.* 1626, 118–135. doi: 10.1016/j.brainres.2015.02.014

Huettig, F., and Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Lang. Cogn. Neurosci.* 31, 80–93. doi: 10.1080/23273798.2015.1047459

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Biling. Lang. Congn.* 15, 422–433. doi: 10.1017/S1366728911000678

Hurley, R., and Bishop, J. (2016). 'Prosodic and individual influences on the interpretation of', in. *Proc. Speech Prosody* 2016, 193–197. doi: 10.21437/SpeechProsody.2016-40

Hurley, R. S. E., Losh, M., Parlier, M., Reznick, J. S., and Piven, J. (2007). The broad autism phenotype questionnaire. *J. Autism Dev. Disord.* 37, 1679–1690. doi: 10.1007/s10803-006-0299-3

James, A. N., Fraundorf, S. H., Lee, E. K., and Watson, D. G. (2018). Individual differences in syntactic processing: is there evidence for reader-text interactions? *J. Mem. Lang.* 102, 155–181. doi: 10.1016/j.jml.2018.05.006

Johnson-Laird, P. N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness. Cambridge: Cambridge University Press.

Johnson, E., and Arnold, J. E. (2021). Individual Differences in Print Exposure Predict Use of Implicit Causality in Pronoun Comprehension and Referential Prediction. *Frontiers in Psychology* 12. doi: 10.3389/fpsyg.2021.672109

Johnson-Laird, P. N., and Garnham, A. (1980). Descriptions and discourse models. *Linguist. Philos.* 3, 371–393. doi: 10.1007/BF00401691

Jouravlev, O., Kell, A. J. E., Mineroff, Z., Haskins, A. J., Ayyash, D., Kanwisher, N., et al. (2020). Reduced language lateralization in autism and the broader autism phenotype as assessed with robust individual-subjects analyses. *Autism Res* 13, 1746–1761. doi: 10.1002/aur.2393

Jun, S.-A., and Bishop, J. (2015a). Priming implicit prosody: prosodic boundaries and individual differences. *Lang. Speech* 58, 459–473. doi: 10.1177/0023830914563368

Jun, S.-A., and Bishop, J. (2015b). "Prominence in relative clause attachment: evidence from prosodic priming" in Explicit and implicit prosody in sentence processing: Studies in honor of Janet dean Fodor. eds. L. Frazier and E. Gibson (Cham: Springer International Publishing), 217–240.

Kamp, H. (1981). "A theory of truth and semantic representation" in Formal semantics - the essential readings. eds. P. Portner and B. H. Partee (Padstow, Cornwall: Blackwell), 189–222.

Karmiloff-Smith, A. (1992). Beyond modularity: A developmental perspective on cognitive science. Cambridge, MA, US: The MIT Press.

Kidd, E., Donnelly, S., and Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends Cogn. Sci.* 22, 154–169. doi: 10.1016/j.tics.2017.11.006

Kintsch, W., and van Dijk, T. A. (1978). Towards a model of discourse comprehension and production. *Psychol. Rev.* 85, 363–394. doi: 10.1037/0033-295X.85.5.363

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *J. Exp. Psychol.* 55, 352–358. doi: 10.1037/h0043688

Kovacs, K., and Conway, A. R. A. (2016). Process overlap theory: a unified account of the general factor of intelligence. *Psychol. Inq.* 27, 151–177. doi: 10.1080/1047840X.2016.1153946

Kulesz, P. A., Francis, D. J., Barnes, M. A., and Fletcher, J. M. (2016). The influence of properties of the test and their interactions with reader characteristics on reading comprehension: an explanatory item response study. *J. Educ. Psychol.* 108, 1078–1097. doi: 10.1037/edu0000126

LaPointe, L. B., and Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 1118–1133. doi: 10.1037/0278-7393.16.6.1118

Langlois, V. J., and Arnold, J. E. (2020). Print exposure explains individual differences in using syntactic but not semantic cues for pronoun comprehension. *Cognition*, 197, Article 104155.

Lascarides, A., and Asher, N. (2007). "Segmented discourse representation theory: dynamic semantics with discourse structure"' in Computing meaning. eds. H. Bunt and R. Muskens (Berlin: Springer Netherlands), 87–124.

Lemhöfer, K., and Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behav. Res. Methods* 44, 325–343. doi: 10.3758/s13428-011-0146-0

Lin, N., Yang, X., Li, J., Wang, S., Hua, H., Ma, Y., et al. (2018). Neural correlates of three cognitive processes involved in theory of mind and discourse comprehension. *Cogn. Affect. Behav. Neurosci.* 18, 273–283. doi: 10.3758/s13415-018-0568-6

Lorenzen, J., and Baumann, S. (2024) 'Does communicative skill predict individual variability in the prosodic encoding of lexical and referential givenness?. Proceedings of Speech Prosody, Leiden.

Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hearing Res.* 50, 940–967. doi: 10.1044/1092-4388(2007/067)

Marsh, H. W., Morin, A. J. S., Parker, P. D., and Kaur, G. (2014). Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annu. Rev. Clin. Psychol.* 10, 85–110. doi: 10.1146/annurev-clinpsy-032813-153700

Martin, J. D., Shipstead, Z., Harrison, T. L., Redick, T. S., Bunting, M., and Engle, R. W. (2020). The role of maintenance and disengagement in predicting reading comprehension and vocabulary learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 46, 140–154. doi: 10.1037/xlm0000705

Martin-Chang, S. L., and Gould, O. N. (2008). Revisiting print exposure: exploring differential links to vocabulary, comprehension and reading rate. *J. Res. Read.* 31, 273–284. doi: 10.1111/j.1467-9817.2008.00371.x

Matthews, D., Biney, H., and Abbot-Smith, K. (2018). Individual differences in Children's pragmatic ability: a review of associations with formal language, social cognition, and executive functions. *Lang. Learn. Dev* 14, 186–223. doi: 10.1080/15475441.2018.1455584

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

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., and Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *J. Exp. Psychol. Gen.* 130, 621–640. doi: 10.1037//0096-3445.130.4.621

Mol, S. E., and Bus, A. G. (2011). To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. *Psychol. Bull.* 137, 267–296. doi: 10.1037/a0021890

Moore, M., and Gordon, P. C. (2015). Reading ability and print exposure: item response theory analysis of the author recognition test. *Behav. Res. Methods* 47, 1095–1109. doi: 10.3758/s13428-014-0534-3

Msika, E.-F., Despres, M., Piolino, P., and Narme, P. (2024). Dynamic and/or multimodal assessments for social cognition in neuropsychology: results from a systematic literature review. *Clin. Neuropsychol.* 38, 922–962. doi: 10.1080/13854046.2023.2266172

Nelson, H. E. (1991). National Adult Reading Test (NART). New York: Springer.

Newen, A., and Vogeley, K. (2003). Self-representation: searching for a neural signature of self-consciousness. *Conscious. Cogn.* 12, 529–543. doi: 10.1016/S1053-8100(03)00080-1

Nieuwland, M. S., Ditman, T., and Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: an ERP investigation of informativeness and pragmatic abilities. *J. Mem. Lang.* 63, 324–346. doi: 10.1016/j.jml.2010.06.005

Nieuwland, M. S., and Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: evidence from ERPs. *Brain Res.* 1118, 155–167. doi: 10.1016/j.brainres.2006.08.022

Norris (2018) Developing C-tests for estimating proficiency in foreign language research. Berlin: Springer.

Oberauer, K., Süß, H. M., Schulze, R., Wilhelm, O., and Wittmann, W. W. (2000). Working memory capacity - facets of a cognitive ability construct. *Personal. Individ. Differ.* 29, 1017–1045. doi: 10.1016/S0191-8869(99)00251-2

Orrico, R., and D'Imperio, M. (2020). 'Individual empathy levels affect gradual intonation-meaning mapping: the case of biased questions in Salerno Italian', *laboratory. Phonology* 11:238. doi: 10.5334/labphon.238

Park, H. I., Solon, M., Dehghan-Chaleshtori, M., and Ghanbar, H. (2022). Proficiency reporting practices in research on second language acquisition: have we made any Progress? *Lang. Learn.* 72, 198–236. doi: 10.1111/lang.12475

Perfetti, C., and Stafura, J. (2013). Word knowledge in a theory of Reading comprehension. *Sci. Stud. Read.* 18, 22–37. doi: 10.1080/10888438.2013.827687

Pineda-Alhucema, W., Aristizabal, E., Escudero-Cabarcas, J., Acosta-López, J. E., and Vélez, J. I. (2018). Executive function and theory of mind in children with ADHD: a systematic review. *Neuropsychol. Rev.* 28, 341–358. doi: 10.1007/s11065-018-9381-9

Repp, S. (2020). The prosody of Wh-exclamatives and Wh-questions in German: speech act differences, information structure, and sex of speaker. *Lang. Speech* 63, 306–361. doi: 10.1177/0023830919846147

Repp, S., and Seeliger, H. (2020). Prosodic prominence in polar questions and Exclamatives. *Front. Commun* 5:53. doi: 10.3389/fcomm.2020.00053

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., and Meints, J. (2009). "Conscientiousness" in Handbook of individual differences in social behavior (New York, NY: The Guilford Press), 369–381.

Roberts, B., Lejuez, C., Krueger, R. F., Richards, J. M., and Hill, P. L. (2012). What is conscientiousness and how can it be assessed? *Dev. Psychol.* 50, 1315–1330. doi: 10.1037/a0031109

Rönnberg, J., Holmer, E., and Rudner, M. (2021). Cognitive hearing science: three memory systems, two approaches, and the ease of language understanding model. *JSLHR* 64, 359–370. doi: 10.1044/2020_JSLHR-20-00007

Rönnberg, J., Holmer, E., and Rudner, M. (2022a). "The ease of language understanding model"in *The Cambridge Handbook of Working Memory and Language*. (Eds.) Schwieter J. W, Wen Z., (Edward), Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Signoret, C., Andin, J., and Holmer, E. (2022b). The cognitive hearing science perspective on perceiving, understanding, and remembering language: the ELU model. *Front. Psychol.* 13:967260. doi: 10.3389/fpsyg.2022.967260

Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Lang. Test.* 15, 1–20. doi: 10.1177/026553229801500101

Rouder, J. N., and Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychon. Bull. Rev.* 26, 452–467. doi: 10.3758/s13423-018-1558-y

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychol. Sci.* 2, 179–183. doi: 10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T., and Babcock, R. (1991). Decomposing adult age differences in working memory. *Dev. Psychol.* 27, 763–776. doi: 10.1037/0012-1649.27.5.763

Scholman, M. C. J., Demberg, V., and Sanders, T. J. M. (2020). Individual differences in expecting coherence relations: exploring the variability in sensitivity to contextual signals in discourse. *Discourse Process.* 61, 381–403. doi: 10.1080/0163853X.2024.2325262

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *J. Educ. Res.* 99, 323–338. doi: 10.3200/JOER.99.6.323-338

Schroeder, S., Segbers, J., and Schröter, P. (2016). Der Kinder-Titelrekognitionstest (K-TRT). *Diagnostica* 62, 16–30. doi: 10.1026/0012-1924/a000131

Seeliger, H., and Repp, S. (2023). 'Information-structural surprises? Contrast, givenness, and (the lack of) accent shift and deaccentuation in non-assertive speech acts', *laboratory. Phonology* 14:6451. doi: 10.16995/labphon.6451

Simon, J. R., Sly, P. E., and Vilapakkam, S. (1981). Effect of compatibility of S-R mapping on reactions toward the stimulus source. *Acta Psychol.* 47, 63–81. doi: 10.1016/0001-6918(81)90039-1

Sörqvist, P., Dahlström, Ö., Karlsson, T., and Rönnberg, J. (2016). Concentration: the neural underpinnings of how cognitive load shields against distraction. *Front. Hum. Neurosci.* 10:221. doi: 10.3389/fnhum.2016.00221

Stanovich, K. E., and West, R. F. (1989). Exposure to print and orthographic processing. *Read. Res. Q.* 24, 402–433. doi: 10.2307/747605

Sternberg, S. (1969). Memory-scanning: mental processes revealed by reaction-time experiments. *Am. Sci.* 57, 421–457

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Tighe, E. L., and Schatschneider, C. (2016). Modeling the relations among morphological awareness dimensions, vocabulary knowledge, and Reading comprehension in adult basic education students. *Front. Psychol.* 7:86. doi: 10.3389/fpsyg.2016.00086

Torregrossa, J., Andreou, M., Bongartz, C., and Tsimpli, I. M. (2021). Bilingual acquisition of reference: the role of language experience, executive functions and cross-linguistic effects. *Biling. Lang. Congn.* 24, 694–706. doi: 10.1017/S1366728920000826

Torregrossa, J., Eisenbeiß, S., and Bongartz, C. (2023). Boosting bilingual metalinguistic awareness under dual language activation: some implications for bilingual education. *Lang. Learn.* 73, 683–722. doi: 10.1111/lang.12552

Torregrossa, J., Listanti, A., Bongartz, C., and Marinis, T. (2024). 'Adding discourse to sentence repetition tasks: under which conditions does bilingual children's performance improve?', *Research methods. Appl. Linguis.* 3:100107. doi: 10.1016/j.rmal.2024.100107

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Stud. Second. Lang. Acquis.* 33, 339–372. doi: 10.1017/S0272263111000015

Tremblay, A., and Garrison, M. D. (2010) Cloze tests: a tool for proficiency assessment in research on L2 French. Selected Proceedings of the 2008 Second Language Research Forum: Exploring SLA Perspectives, Positions, and Practices, 73-88, Cascadilla Proceedings Project, Somerville, MA, USA.

Truscott, J., and Smith, M. S. (2022). "Working memory in the modular cognition framework" in *The Cambridge Handbook of Working Memory and Language*. (Eds.) Schwieter J. W, Wen Z., (Edward), Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

Turner, M. L., and Engle, R. W. (1989). Is working memory capacity task dependent? *J. Mem. Lang.* 28, 127–154. doi: 10.1016/0749-596X(89)90040-5

van den Noort, M., Bosch, P., Haverkort, M., and Hugdahl, K. (2008). A standard computerized version of the Reading span test in different languages. *Eur. J. Psychol. Assess.* 24, 35–42. doi: 10.1027/1015-5759.24.1.35

Vogeley, K., and Bartels, A. (2011). The explanatory value of representations in cognitive neuroscience. Stanford & Mentis-Verlag: CSLI Publications.

Vogelzang, M., Guasti, M. T., van Rijn, H., and Hendriks, P. (2021). How children process reduced forms: a computational cognitive modeling approach to pronoun processing in discourse. *Cogn. Sci.* 45:e12951. doi: 10.1111/cogs.12951

Vogelzang, M., and Torregrossa, J. (2024). Identifying bilingual reference profiles: a cluster-analysis approach to reference production among Greek–Italian bilingual children. *Lang. Cogn.* 16, 1–24. doi: 10.1017/langcog.2024.48

von Heusinger, K., and Schumacher, P. B. (2019). Discourse prominence: definition and application. *J. Pragmat.* 154, 117–127. doi: 10.1016/j.pragma.2019.07.025

Wehrle, S. (2023). Conversation and intonation in autism: A multi-dimensional analysis. Berlin: Language Science Press.

Wehrle, S., Cangemi, F., Vogeley, K., and Grice, M. (2022). New evidence for melodic speech in autism Spectrum disorder. *Proc. Speech Prosody* 2022, 37–41. doi: 10.21437/SpeechProsody.2022-8

Wehrle, S., Grice, M., and Vogeley, K. (2024a). Filled pauses produced by autistic adults differ in prosodic realisation, but not rate or lexical type. *J. Autism Dev. Disord.* 54, 2513–2525. doi: 10.1007/s10803-023-06000-y

Wehrle, S., Vogeley, K., and Grice, M. (2024b). Backchannels in conversations between autistic adults are less frequent and less diverse prosodically and lexically. *Lang. Cogn.* 16, 108–133. doi: 10.1017/langcog.2023.21

Wetzel, M., Zufferey, S., and Gygax, P. (2020). Second language acquisition and the mastery of discourse connectives: assessing the factors that hinder L2-learners from mastering French connectives. *Languages* 5:35. doi: 10.3390/languages5030035

Whitney, P., Arnett, P. A., Driver, A., and Budd, D. (2001). Measuring central executive functioning: what's in a reading span? *Brain Cogn.* 45, 1–14. doi: 10.1006/brcg.2000.1243

Wilhelm, O., Hildebrandt, A. H., and Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Front. Psychol.* 4:433. doi: 10.3389/fpsyg.2013.00433

Yap, M. J., Balota, D. A., Sibley, D. E., and Ratcliff, R. (2011). 'Individual differences in visual word recognition: insights from the English lexicon project', *journal of experimental psychology. Human Percep. Perform.* 38:53.

Yu, A. C. L., and Zellou, G. (2019). Individual differences in language processing: phonology. *Annu. Rev. Linguist.* 5, 131–150. doi: 10.1146/annurev-linguistics-011516-033815

Zimmermann, J. T., Meuser, S., Hinterwimmer, S., and Vogeley, K. (2021). Preserved perspective taking in free indirect discourse in autism Spectrum disorder. *Front. Psychol.* 12:675633. doi: 10.3389/fpsyg.2021.675633

Zufferey, S., and Gygax, P. (2020). "Roger broke his tooth. However, he went to the dentist": why some readers struggle to evaluate wrong (and right) uses of connectives. *Discourse Process.* 57, 184–200. doi: 10.1080/0163853X.2019.1607446