Differences in linguistic complexity between experimental conditions generally lead to **increased cognitive load** for participants and can **negatively impact response validity**, especially for individuals with lower cognitive abilities or in more demanding tasks.

1. Introduction

The relationship between linguistic complexity, cognitive load, and response validity is a central concern in psychological research, particularly in the design and interpretation of experimental tasks and assessments. Numerous studies have demonstrated that increasing linguistic complexity—through longer sentences, more abstract vocabulary, or complex syntactic structures—tends to elevate cognitive load, as measured by both subjective and objective indicators such as response times, self-reported effort, and physiological measures like EEG and pupil dilation (Cohen et al., 2021; Castro-Meneses et al., 2019; Demberg & Sayeed, 2016; Just et al., 1996; Vogelzang et al., 2020). This heightened cognitive load can, in turn, affect the validity of participant responses, with evidence suggesting that individuals with lower cognitive abilities or higher anxiety are disproportionately affected by complex language, leading to slower responses and potentially less accurate or less valid data (Cohen et al., 2021; Güvendir & Uzun, 2023; Sung et al., 2020; Révész et al., 2015). Experimental manipulations of linguistic complexity have shown robust effects on cognitive processing across modalities (e.g., reading, listening, writing) and populations (e.g., native speakers, L2 learners, individuals with cognitive impairments) (Lee, 2018; Révész et al., 2017; Brunfaut & Révész, 2015; Shain et al., 2021; Sung et al., 2020; Awwad & Tavakoli, 2019; Xu et al., 2021; Cho, 2018; Kim et al., 2015). However, the impact on response validity is nuanced, with some studies indicating that increased complexity primarily affects processing speed and effort rather than the accuracy of responses, while others highlight the risk of reduced data quality in more demanding conditions (Cohen et al., 2021; Révész et al., 2015; Embretson, 2023; Hunter, 2021). The literature also underscores the importance of considering individual differences (e.g., working memory, language proficiency, anxiety) and task characteristics (e.g., modality, support, genre) when evaluating the effects of linguistic complexity on cognitive load and response validity (Révész et al., 2017; Awwad & Tavakoli, 2019; Güvendir & Uzun, 2023; Cho, 2018; Kim et al., 2015; Yoon, 2021; Li et al., 2024). Overall, the evidence supports a strong link between linguistic complexity, cognitive load, and the potential for compromised response validity, emphasizing the need for careful task design in psychological research.

2. Methods

A comprehensive literature review was conducted using Consensus, which aggregates over 170 million research papers from sources including Semantic Scholar and PubMed. The search strategy involved 21 targeted queries across 8 search groups, focusing on linguistic complexity, cognitive load, and response validity in psychological research. In total, 1,051 papers were identified, 592 were screened, 379 were deemed eligible, and the top 50 most relevant papers were included in this review.

Search Strategy

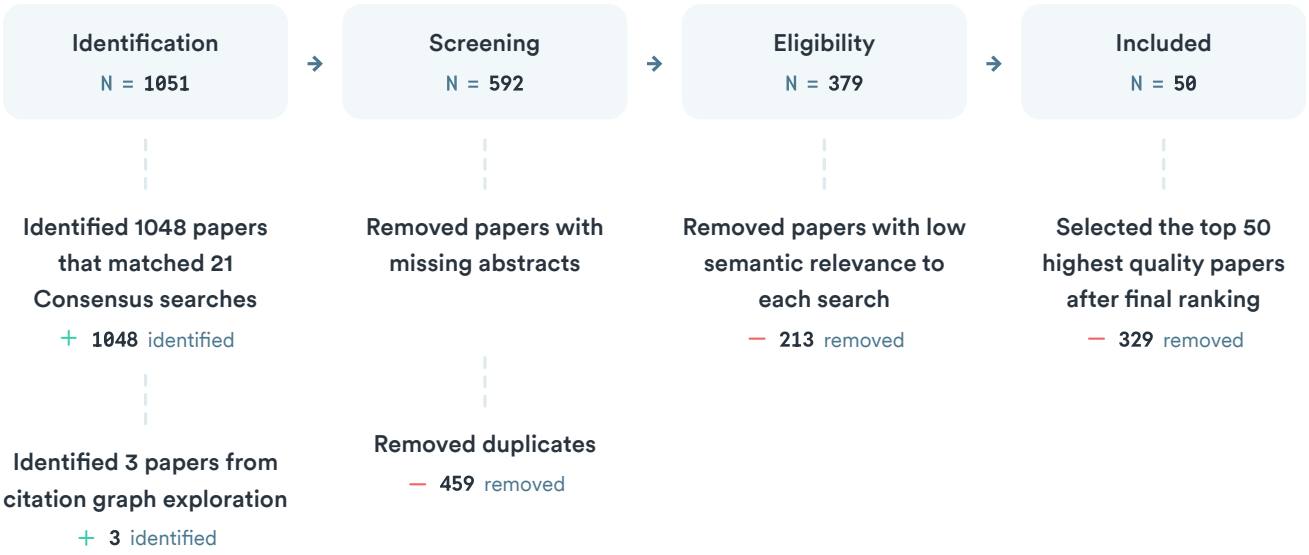

FIGURE 1 Flow diagram of the literature search and selection process.

Eight unique search strategies were used, systematically targeting foundational theories, experimental manipulations, interdisciplinary perspectives, and methodological critiques.

3. Results

3.1. Effects of Linguistic Complexity on Cognitive Load

Multiple studies confirm that increasing linguistic complexity—such as longer sentences, more abstract words, or complex syntax—leads to higher cognitive load, as evidenced by longer response times, increased self-reported effort, and physiological markers (e.g., EEG theta power, pupil dilation) (Cohen et al., 2021; Castro-Meneses et al., 2019; Demberg & Sayeed, 2016; Just et al., 1996; Vogelzang et al., 2020). These effects are robust across modalities (reading, listening, writing) and are particularly pronounced in individuals with lower cognitive abilities or higher anxiety (Cohen et al., 2021; Güvendir & Uzun, 2023; Sung et al., 2020; Révész et al., 2015; Shain et al., 2021; Awwad & Tavakoli, 2019; Cho, 2018; Kim et al., 2015; Mavrou, 2020).

3.2. Impact on Response Validity

While increased cognitive load due to linguistic complexity often results in slower responses, the impact on response validity (i.e., the accuracy and interpretability of responses) is more nuanced. Some studies report that validity is not necessarily compromised within plausible response ranges, but others find that data quality can suffer, especially for participants with lower cognitive resources or in high-demand tasks (Cohen et al., 2021; Révész et al., 2015; Embretson, 2023; Hunter, 2021; Yoon, 2021). The risk of invalid or less reliable responses increases as cognitive demands outpace participant capacity.

3.3. Individual Differences and Moderators

Individual factors such as working memory, language proficiency, and anxiety significantly moderate the effects of linguistic complexity on cognitive load and response validity (Révész et al., 2017; Awwad & Tavakoli, 2019; Güvendir & Uzun, 2023; Cho, 2018; Kim et al., 2015; Li et al., 2024; Mavrou, 2020). For example, high working memory capacity can buffer the negative effects of complex language, while high anxiety or low proficiency can exacerbate them.

3.4. Measurement and Methodological Considerations

Subjective cognitive load questionnaires are generally reliable, but their validity can be affected by translation, item formulation, and participant understanding (Krieglstein et al., 2022; Krieglstein et al., 2023). Objective measures (e.g., EEG, pupil dilation) provide convergent evidence but may differ in sensitivity to complexity manipulations (Castro-Meneses et al., 2019; Demberg & Sayeed, 2016). Dual-task paradigms and response time modeling are effective for validating cognitive load manipulations and assessing their impact on response processes (Lee, 2018; Dutilh et al., 2018; Révész et al., 2015; Engonopoulos, 2013; Embretson, 2023; Hunter, 2021).

Key Papers

Paper	Methodology	Population	Main Findings	Cognitive Load Measure
(Cohen et al., 2021)	Mixed models, psycholinguistic analysis	581 adults with neurological disorders	Longer, less imageable items increased response time, especially for lower cognitive ability	Response time
(Castro-Meneses et al., 2019)	Experimental, EEG	35 postgraduate L2 students	Higher linguistic complexity increased subjective and EEG-based cognitive load	EEG theta power, self-report
(Révész et al., 2017)	Keystroke logging, recall	73 advanced L2 writers	Content support reduced pausing and increased linguistic complexity	Keystroke data, recall
(Révész et al., 2015)	Dual-task, self-ratings, expert judgment	96 students, 61 ESL teachers	More complex tasks increased cognitive effort on most measures	Dual-task, self-report, expert rating
(Just et al., 1996)	fMRI	Not specified	Greater linguistic complexity increased neural activation in language areas	fMRI (brain activation)

FIGURE 2 Comparison of key studies on linguistic complexity, cognitive load, and response validity.

Top Contributors

Type	Name	Papers
Author	A. Révész	(Révész et al., 2017; Brunfaut & Révész, 2015; Révész et al., 2015)
Author	Felix Krieglstein	(Krieglstein et al., 2022; Krieglstein et al., 2023)
Author	Jiyong Lee	(Lee, 2018)
Institution	Indiana University	(Hunter, 2021; Just et al., 1996)
Institution	University of Cambridge	(Révész et al., 2017; Brunfaut & Révész, 2015; Révész et al., 2015)
Institution	University of Hagen	(Krieglstein et al., 2022; Krieglstein et al., 2023)

FIGURE 3 Authors & institutions that appeared most frequently in the included papers.

4. Discussion

The literature provides strong evidence that linguistic complexity is a significant driver of cognitive load in psychological research tasks, with clear implications for response validity. The most robust findings indicate that increased complexity leads to longer response times, higher self-reported effort, and greater neural activation, especially in language-selective brain regions (Cohen et al., 2021; Castro-Meneses et al., 2019; Demberg & Sayeed, 2016; Just et al., 1996; Vogelzang et al., 2020). While some studies suggest that response validity (accuracy) may be preserved within certain bounds, others highlight the risk of compromised data quality, particularly for participants with lower cognitive resources or in high-demand conditions (Cohen et al., 2021; Révész et al., 2015; Embretson, 2023; Hunter, 2021; Yoon, 2021). The reliability of subjective cognitive load measures is generally high, but their validity can be affected by translation and item formulation, underscoring the need for careful instrument design and validation (Krieglstein et al., 2022; Krieglstein et al., 2023). Objective measures such as EEG and pupil dilation offer valuable convergent evidence but may differ in sensitivity and practicality (Castro-Meneses et al., 2019; Demberg & Sayeed, 2016). Individual differences—such as working memory, language proficiency, and anxiety—play a critical role in moderating the effects of linguistic complexity, suggesting that task design should account for participant variability to ensure valid and interpretable results (Révész et al., 2017; Awwad & Tavakoli, 2019; Güvendir & Uzun, 2023; Cho, 2018; Kim et al., 2015; Li et al., 2024; Mavrou, 2020). Overall, the research underscores the importance of aligning linguistic complexity with participant abilities and research goals to optimize both cognitive load and response validity.

Claims and Evidence Table

Claim	Evidence Strength	Reasoning	Papers
Increasing linguistic complexity increases cognitive load in experimental tasks	Strong	Supported by convergent evidence from response times, self-report, EEG, and fMRI across multiple populations and modalities	(Cohen et al., 2021; Castro-Meneses et al., 2019; Demberg & Sayeed, 2016; Just et al., 1996; Vogelzang et al., 2020)
Higher cognitive load can reduce response validity, especially for low-ability participants	Strong	Multiple studies show slower, less accurate, or less reliable responses under high complexity, particularly in vulnerable groups	(Cohen et al., 2021; Révész et al., 2015; Güvendir & Uzun, 2023; Sung et al., 2020; Embretson, 2023; Hunter, 2021)
Subjective cognitive load questionnaires are reliable but may have validity issues in translation or with children	Moderate	Meta-analyses and validation studies highlight reliability but caution about construct validity and translation effects	(Krieglstein et al., 2022; Krieglstein et al., 2023)
Individual differences (working memory, proficiency, anxiety) moderate the impact of complexity on cognitive load and validity	Moderate	Consistent findings across studies show that these factors buffer or exacerbate effects	(Révész et al., 2017; Awwad & Tavakoli, 2019; Güvendir & Uzun, 2023; Cho, 2018; Kim et al., 2015; Li et al., 2024; Mavrou, 2020)
Increased complexity affects processing speed more than response accuracy in some contexts	Moderate	Some studies find that validity is preserved within plausible response ranges, but speed and effort are affected	(Cohen et al., 2021; Hunter, 2021; Yoon, 2021)
Effects of linguistic complexity on cognitive load are less pronounced in highly proficient or motivated participants	Moderate	Evidence suggests that high proficiency or motivation can mitigate negative effects	(Révész et al., 2017; Awwad & Tavakoli, 2019; Li et al., 2024)

FIGURE 4 Key claims and support evidence identified in these papers.

5. Conclusion

In summary, the literature strongly supports that differences in linguistic complexity between experimental conditions increase participant cognitive load and can negatively impact response validity, particularly for individuals with lower cognitive abilities or in more demanding tasks. The effects are robust across modalities and populations, but are moderated by individual differences and task characteristics. Careful task and instrument design, as well as consideration of participant variability, are essential for optimizing both cognitive load and response validity in psychological research.

5.1. Research Gaps

Despite substantial evidence, gaps remain in understanding the nuanced effects of linguistic complexity on response validity across diverse populations, task types, and real-world settings. There is also a need for more research on the interplay between cognitive load, motivation, and affect, as well as the development of more sensitive and universally valid measurement tools.

Research Gaps Matrix

Topic / Attribute	Native Speakers	L2 Learners	Low Working Memory	High Working Memory	Children
Reading tasks	7	5	2	3	1
Listening tasks	6	4	2	2	1
Writing tasks	4	6	1	2	GAP
Dual-task paradigms	3	2	1	1	GAP
Neuroimaging	2	1	GAP	1	GAP

FIGURE 5 Matrix showing research coverage by topic and population/attribute.

5.2. Open Research Questions

Future research should further explore the mechanisms by which linguistic complexity affects response validity, develop more sensitive and universally valid measurement tools, and investigate interventions to mitigate negative effects for vulnerable populations.

Question	Why
How does linguistic complexity affect response validity in real-world, high-stakes psychological assessments?	Real-world settings may introduce additional variables and higher stakes, potentially amplifying or mitigating the effects observed in laboratory studies.
What interventions can reduce the negative impact of linguistic complexity on cognitive load for low-ability or high-anxiety participants?	Identifying effective interventions could improve data quality and fairness in psychological research and assessment.
How do affective and motivational factors interact with cognitive load and linguistic complexity to influence response validity?	Understanding these interactions could inform more holistic models of task performance and lead to better task design.

FIGURE 6 Open research questions and their significance for future studies.

In conclusion, aligning linguistic complexity with participant abilities and research goals is crucial for optimizing cognitive load and ensuring valid, reliable responses in psychological research.

These papers were sourced and synthesized using Consensus, an AI-powered search engine for research. Try it at https://consensus.app

References

Cohen, M., Boulton, A., Lanzi, A., Sutherland, E., & Pompon, R. (2021). Psycholinguistic features, design attributes, and respondent-reported cognition predict response time to patient-reported outcome measure items. *Quality of Life Research*, 30, 1693 - 1704. https://doi.org/10.1007/s11136-021-02778-5

Lee, J. (2018). Task Complexity, Cognitive Load, and L1 Speech. *Applied Linguistics*, 40, 506-539. https://doi.org/10.1093/APPLIN/AMX054

Castro-Meneses, L., Kruger, J., & Doherty, S. (2019). Validating theta power as an objective measure of cognitive load in educational video. *Educational Technology Research and Development*, 68, 181 - 202. https://doi.org/10.1007/s11423-019-09681-4

Krieglstein, F., Beege, M., Rey, G., Ginns, P., Krell, M., & Schneider, S. (2022). A Systematic Meta-analysis of the Reliability and Validity of Subjective Cognitive Load Questionnaires in Experimental Multimedia Learning Research. *Educational Psychology Review*, 34, 2485 - 2541. https://doi.org/10.1007/s10648-022-09683-4

Krieglstein, F., Beege, M., Rey, G., Sanchez-Stockhammer, C., & Schneider, S. (2023). Development and Validation of a Theory-Based Questionnaire to Measure Different Types of Cognitive Load. *Educational Psychology Review*, 35, 1-37. https://doi.org/10.1007/s10648-023-09738-0

Dutilh, G., Annis, J., Brown, S., Cassey, P., Evans, N., Grasman, R., Hawkins, G., Heathcote, A., Holmes, W., Krypotos, A., Kupitz, C., Leite, F., Lerche, V., Lin, Y., Logan, G., Palmeri, T., Starns, J., Trueblood, J., Van Maanen, L., Van Ravenzwaaij, D., Vandekerckhove, J., Visser, I., Voss, A., White, C., Wiecki, T., Rieskamp, J., & Donkin, C. (2018). The Quality of Response Time Data Inference: A Blinded, Collaborative Assessment of the Validity of Cognitive Models. *Psychonomic Bulletin & Review*, 26, 1051 - 1069. https://doi.org/10.3758/s13423-017-1417-2

Demberg, V., & Sayeed, A. (2016). The Frequency of Rapid Pupil Dilations as a Measure of Linguistic Processing Difficulty. *PLoS ONE*, 11. https://doi.org/10.1371/journal.pone.0146194

Révész, A., Kourtali, N., & Mazgutova, D. (2017). Effects of Task Complexity on L2 Writing Behaviors and Linguistic Complexity. *Language Learning*, 67, 208-241. https://doi.org/10.1111/LANG.12205

Brunfaut, T., & Révész, A. (2015). The Role of Task and Listener Characteristics in Second Language Listening.. *TESOL Quarterly*, 49, 141-168. https://doi.org/10.1002/TESQ.168

Révész, A., Michel, M., & Gilabert, R. (2015). MEASURING COGNITIVE TASK DEMANDS USING DUAL-TASK METHODOLOGY, SUBJECTIVE SELF-RATINGS, AND EXPERT JUDGMENTS. *Studies in Second Language Acquisition*, 38, 703 - 737. https://doi.org/10.1017/S0272263115000339

Shain, C., Blank, I., Fedorenko, E., Gibson, E., & Schuler, W. (2021). Robust Effects of Working Memory Demand during Naturalistic Language Comprehension in Language-Selective Cortex. *The Journal of Neuroscience*, 42, 7412 - 7430. https://doi.org/10.1101/2021.09.18.460917

Sung, J., Choi, S., Eom, B., Yoo, J., & Jeong, J. (2020). Syntactic Complexity as a Linguistic Marker to Differentiate Mild Cognitive Impairment From Normal Aging.. *Journal of speech, language, and hearing research : JSLHR*, 1-14. https://doi.org/10.1044/2020_JSLHR-19-00335

Awwad, A., & Tavakoli, P. (2019). Task complexity, language proficiency and working memory: Interaction effects on second language speech performance. *International Review of Applied Linguistics in Language Teaching*, 60, 169 - 196. https://doi.org/10.1515/IRAL-2018-0378

Engonopoulos, N. (2013). Linguistic Complexity and Cognitive Load in a Dual-Task Context. **.

Güvendir, E., & Uzun, K. (2023). L2 writing anxiety, working memory, and task complexity in L2 written performance. *Journal of Second Language Writing*. https://doi.org/10.1016/j.jslw.2023.101016

Embretson, S. (2023). Understanding examinees' item responses through cognitive modeling of response accuracy and response times. *Large-scale Assessments in Education*, 11, 1-16. https://doi.org/10.1186/s40536-023-00158-8

Xu, T., Zhang, L., & Gaffney, J. (2021). EXAMINING THE RELATIVE EFFECTS OF TASK COMPLEXITY AND COGNITIVE DEMANDS ON STUDENTS' WRITING IN A SECOND LANGUAGE. *Studies in Second Language Acquisition*, 44, 483 - 506. https://doi.org/10.1017/S0272263121000310

Cho, M. (2018). Task complexity, modality, and working memory in L2 task performance. *System*, 72, 85-98. https://doi.org/10.1016/J.SYSTEM.2017.10.010

Kim, Y., Payant, C., & Pearson, P. (2015). THE INTERSECTION OF TASK-BASED INTERACTION, TASK COMPLEXITY, AND WORKING MEMORY. *Studies in Second Language Acquisition*, 37, 549 - 581. https://doi.org/10.1017/S0272263114000618

Hunter, C. (2021). Dual-Task Accuracy and Response Time Index Effects of Spoken Sentence Predictability and Cognitive Load on Listening Effort. *Trends in Hearing*, 25. https://doi.org/10.1177/23312165211018092

Yoon, H. (2021). Challenging the connection between task perceptions and language use in L2 writing: Genre, cognitive task complexity, and linguistic complexity. *Journal of Second Language Writing*. https://doi.org/10.1016/j.jslw.2021.100857

Li, C., Wei, L., & Lu, X. (2024). Task complexity and L2 writing performance of young learners: Contributions of cognitive and affective factors. *The Modern Language Journal*. https://doi.org/10.1111/modl.12954

Mavrou, I. (2020). Working memory, executive functions, and emotional intelligence in second language writing. *Journal of Second Language Writing*. https://doi.org/10.1016/j.jslw.2020.100758

Just, M., Carpenter, P., Keller, T., Eddy, W., & Thulborn, K. (1996). Brain Activation Modulated by Sentence Comprehension. *Science*, 274, 114 - 116. https://doi.org/10.1126/science.274.5284.114

Vogelzang, M., Thiel, C., Rosemann, S., Rieger, J., & Ruigendijk, E. (2020). Neural Mechanisms Underlying the Processing of Complex Sentences: An fMRI Study. *Neurobiology of Language*, 1, 226 - 248. https://doi.org/10.1162/nol_a_00011