ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/378211669

Cognitive components of aging-related increase in word-finding difficulty

Article *in* Aging Neuropsychology and Cognition · February 2024

DOI: 10.1080/13825585.2024.2315774

CITATIONS
5

READS
213

7 authors, including:

Hsi Tiana Wei
McGill University
8 PUBLICATIONS **38** CITATIONS
SEE PROFILE

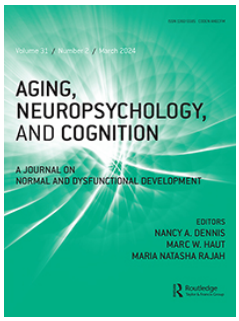
You Zhi Hu
University of Toronto
10 PUBLICATIONS **45** CITATIONS
SEE PROFILE

Mark Chignell
University of Toronto
448 PUBLICATIONS **5,983** CITATIONS
SEE PROFILE

Jed A Meltzer
Baycrest
109 PUBLICATIONS **3,179** CITATIONS
SEE PROFILE

All content following this page was uploaded by Hsi Tiana Wei on 15 February 2024.

The user has requested enhancement of the downloaded file.



Routledge
Taylor & Francis Group

Aging, Neuropsychology, and Cognition

A Journal on Normal and Dysfunctional Development

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/nanc20

Cognitive components of aging-related increase in word-finding difficulty

Hsi T. Wei, Dana Kulzhabayeva, Lella Erceg, Jessica Robin, You Zhi Hu, Mark Chignell & Jed A. Meltzer

To cite this article: Hsi T. Wei, Dana Kulzhabayeva, Lella Erceg, Jessica Robin, You Zhi Hu, Mark Chignell & Jed A. Meltzer (14 Feb 2024): Cognitive components of aging-related increase in word-finding difficulty, Aging, Neuropsychology, and Cognition, DOI: 10.1080/13825585.2024.2315774

To link to this article: https://doi.org/10.1080/13825585.2024.2315774

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 14 Feb 2024.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=nanc20

AGING, NEUROPSYCHOLOGY, AND COGNITION
https://doi.org/10.1080/13825585.2024.2315774

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Cognitive components of aging-related increase in word-finding difficulty

Hsi T. Wei [a,b], Dana Kulzhabayeva[a,b], Lella Erceg[b,c], Jessica Robin[d], You Zhi Hu[e], Mark Chignell[e] and Jed A. Meltzer [a,b]

[a]Department of Psychology, University of Toronto, Toronto, Canada; [b]Rotman Research Institute, Baycrest Hospital, Toronto, Canada; [c]Department of Cell and Systems Biology, University of Toronto, Toronto, Canada; [d]Winterlight Labs, Toronto, Canada; [e]Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada

ABSTRACT

Word-finding difficulty (WFD) is a common cognitive complaint in aging, manifesting both in natural speech and in controlled laboratory tests. Various theories of cognitive aging have addressed WFD, and understanding its underlying mechanisms can help to clarify whether it has diagnostic value for neurodegenerative disease. Two influential "information-universal" theories attribute it to rather broad changes in cognition. The processing speed theory posits a general slowdown of all cognitive processes, while the inhibitory deficit hypothesis (IDH) predicts a specific problem in suppressing irrelevant information. One "information specific" theory of language production, the transmission deficit hypothesis (TDH), posits a breakdown in retrieval of phonological word forms from a corresponding lemma. To adjudicate between these accounts, we administered an online gamified covert naming task featuring picture-word interference (PWI), previously validated to elicit similar semantic interference and phonological facilitation effects as overt naming tasks. 125 healthy adults aged 18 to 85 completed the task, along with a battery of executive function tasks and a naturalistic speech sample to quantify WFD in connected speech. PWI effects provided strong support for the TDH but limited support for IDH, in that semantic interference increased and phonological facilitation decreased across the lifespan. However, neither of these effects on single-word retrieval associated with WFD measured in connected speech. Rather, overall reaction time for word retrieval (controlling for psychomotor slowing) was the best predictor of spontaneous WFD and executive function decline, suggesting processing speed as the key factor, and that verbal reaction time may be an important clinical measure.

ARTICLE HISTORY
Received 14 August 2023
Accepted 17 November 2023

KEYWORDS
Word-finding difficulty; aging; processing speed theory; inhibition deficit hypothesis; transmission deficit hypothesis

1. Introduction

As one of the most common cognitive complaints associated with aging (R. Brown & McNeill, 1966; Cohen, 1994; James & Burke, 2000), word-finding difficulty (WFD) has

CONTACT Hsi T. Wei twei@research.baycrest.org Department of Psychology, University of Toronto, 100 St George St Toronto, ON M5S 3G3, Canada

Supplemental data for this article can be accessed online at https://doi.org/10.1080/13825585.2024.2315774

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

a ubiquitous presence in neurodegenerative pathologies such as frontotemporal dementia and Alzheimer's Disease (AD) (Faber-Langendoen et al., 1988; Kirshner, 2012; Rohrer et al., 2008). Therefore, the severity of WFD may hold diagnostic value in distinguishing healthy aging from incipient neuropathology. To better understand the cognitive components behind age-related WFD, an extensive amount of research on word production in aging has contributed to the development of two distinct viewpoints of cognitive aging – information-universal versus information-specific theories. One main difference between the two approaches in explaining WFD concerns whether the age-related decline in word production arises from deficits specific to the domain of language processing (information-specific) or a more global decline that manifests in many different cognitive domains (information-universal).

1.1. Processing speed theory

One of the earliest information-universal theories advocates decreased processing speed as the common factor that contributes to most of the age-related changes in fluid intelligence, impacting performance on many cognitive tasks (T. Salthouse, 1996). Cognitive slowing is assumed to lower task performance through either time limitation, where the required mental operations cannot be executed within the limited time given, or through a simultaneity failure, where required information is no longer available by the time the mental operation is ready to complete (T. Salthouse, 1996). One finding that is interpreted as supporting processing speed theory is that, when factoring out the general effect of processing speed, age-related differences in more specific cognitive measures (e.g., attention, perception, memory) tend to be removed or at least attenuated (T. Salthouse, 1996; T. A. Salthouse, 2001, 2017). Nonetheless, some domain specific slowing has been observed; for example, age related slowing has been found to be stronger for non-lexical tasks compared to lexical tasks (Hale & Myerson, 1996; Lima et al., 1991). It is likely that a process-specific slowing can exist as an additional layer that modulates, but does not negate, the generalized slowing mechanism (Madden & Allen, 2015; T. Salthouse, 1996). It is clear that older adults are significantly slower than younger adults in completing various cognitive tasks, including word-production tasks such as picture naming, answering questions, or reading written words (see Amrhein, 1995 for a review). In natural speech, older adults also tend to produce more dysfluencies such as unfilled and filled pauses (e.g., "uh" and "um") in between speech and have a generally slower speech rate (Bortfeld et al., 2001; Dennis & Hess, 2016; Horton et al., 2010). Although age is not always a strong predictor of speech disfluencies (e.g., pauses), it does lead to changes in speech characteristics such as speech rate and lexical complexity that in turn predict the production of disfluencies longitudinally (Beier et al., 2023). Explained within the processing speed theory, the word-finding struggles in retrieval tasks and natural discourse likely result from general slowing in activating the mental subprocesses required for word retrieval.

1.2. Inhibition deficit hypothesis

The inhibitory deficit hypothesis (IDH) is an alternative information-universal account that focuses on an age-related decrease in the inhibitory mechanism that filters out unwanted distractions which compete for limited attentional resources (Hasher &

Zacks, 1988). It has garnered support in various cognitive domains, such as visual attention, executive function, and language processing, showing how older adults' decreased ability to ignore distraction can either improve or hinder their task performance depending on the goal of the task. For tasks where attending to distractions leads to better task performance, older adults showed greater benefits than younger adults (Kim et al., 2007; May, 1999; Rowe et al., 2006). For other tasks where distractions hinder performance, older adults exhibited greater difficulties than younger adults (Arbuckle et al., 2000; Pichora-Fuller et al., 1995; Sommers, 1996; Sommers & Danielson, 1999; Tun et al., 2002). Pertaining to the discussion of word production in aging, the IDH postulates that older adults fail to recall names because of their inability to suppress irrelevant associations or competing words during retrieval (Zacks & Hasher, 1994). One piece of evidence against the IDH account, however, is that older adults report fewer alternative words coming to mind when experiencing a tip-of-the-tongue state (TOT), a subtype of WFD in which a person experiences a "feeling of knowing" a desired word without being able to retrieve it (D. M. Burke et al., 1991). Additionally, language comprehension, unlike language production, is well preserved during the aging process (Thornton & Light, 2006), seemingly inconsistent with the information universal framework.

1.3. Transmission deficit hypothesis

In contrast, the information-specific account of aging includes theories that suggest specific cognitive aging mechanisms associated with different types and structures of information (D. M. Burke et al., 1997). The transmission deficit hypothesis (TDH) is a major information-specific theory, providing an explanation for the asymmetrical decline of language production versus comprehension during aging (D. Burke et al., 2012; MacKay & Burke, 1990). The TDH is derived from Node Structure Theory (NST) (Mackay, 1988), which describes the word production process as a top-down activation structure. The top-level nodes represent the sentential system, which contains semantic and syntactic representations. The lower-level nodes reference the phonological system, which includes the phonemes and syllables that comprise a word. The nodes within the sentential system have a many-to-one connection between the propositional nodes (relevant concepts) and the lexical node, resulting in a rather robust top-down activation since excitation of a lexical node can be received through multiple propositions. However, the top-down activation from the sentential to the phonological system relies on a one-to-one connection between the activated lexical representation and its phonological representations, rendering this transmission vulnerable. Importantly, this one-to-one vulnerability only exists during a top-down diverging activation (language production) but not a bottom-up converging activation (language comprehension), explaining the asymmetrical age decline between language production and comprehension (D. Burke et al., 2012; MacKay & Burke, 1990). Based on NST, the transmission deficit hypothesis postulates that, aside from the frequency and recency of word usage, age is another factor that weakens the connections between nodes, causing increased deficits in the vulnerable transmission from lexical to phonological activation of the top-down structure activation, and thereby increasing WFD (MacKay & Burke, 1990).

1.4. Picture-word interference paradigm

As depicted in NST, picture naming can theoretically be segmented into a series of cognitive processing stages (Indefrey, 2011; Indefrey & Levelt, 2004; Levelt et al., 1999; see also Dell et al., 1997; Hillis, 2013). Overall, in the influential language production models such as the Levelt model, the WEAVER++ model (Roelofs, 2003, 2008), there is a consensus that picture naming entails visual recognition and selection of the lexical concept followed by selection of the lemma, retrieval and encoding of the phonological word form, and phonetic encoding and articulation of the picture name. Thus, it is possible to distinguish inhibition and transmission deficit hypotheses-related effects at different stages of a picture-naming task; while the overall RT of word retrieval can reveal cognitive processing speed. Picture-word interference (PWI) is an ideal paradigm to probe different stages of the picture naming process. By manipulating the relationship, and the stimulus onset asynchrony (SOA), between the target picture and distractor word in a PWI paradigm, picture naming reaction time (RT) and accuracy can be influenced in different directions. When the distractor word is categorically-related to the target picture (e.g., *cat* when the target is *dog*) and the distractor word coincides or precedes the picture by a few hundred milliseconds, the picture naming process is usually slowed by the competing word (Abel et al., 2009; Alario et al., 2000; Schriefers et al., 1990; Wei et al., 2022) during the lexical selection stage (Abdel Rahman & Melinger, 2011; Belke et al., 2005; Damian et al., 2001; Schnur et al., 2006). This slowing effect is termed categorical interference (CI). On the other hand, when the distractor word is phonologically-related to the target picture (e.g., *fog* when the target is *dog*) and the distractor word coincides or lags the picture by a few hundred milliseconds, the picture naming process is usually sped up by a priming effect (Abel et al., 2009; Jescheniak & Schriefers, 2001; Schriefers et al., 1990) at the stage of phonological activation (de Zubicaray & McMahon, 2009; de Zubicaray et al., 2002; Meyer, 1996; Meyer & Schriefers, 1991; Roelofs, 1992; Schriefers et al., 1990). This speeding effect is termed phonological facilitation (PF).

Wei et al. (2022) designed a PWI paradigm with covert naming (ending sound judgment with button press) and its corresponding gamified version. They observed that both the experimental and the gamified versions of the PWI covert naming task successfully induced the CI effect at an earlier picture-word SOA and the PF effect at a later SOA. The latency and direction of the CI and PF effects replicated previous findings using the traditional overt naming PWI task (e.g., see review in Abel et al., 2009). The results suggested that even when participants covertly named pictures and judged the ending sound of the picture names, they went through the two main mental stages of word retrieval expected for an overt naming task – lexical selection and phonological activation. The distinct optimal SOAs also supported the proposition that CI happens at an earlier stage of word retrieval during lexical selection (distractor word precedes target picture by 200 ms), while PF happens at a later stage during phonological activation (distractor word coincides with target picture). The ending phoneme was chosen as the judgment target since the ending sound by itself (the coda of a the last syllable), especially without the preceding nucleus/vowel of the syllable, has been found to have no priming effect to target names (e.g., Dumay et al., 2001). Moreover, evidence from PWI phonological priming tasks showed that ending syllable overlap would only elicit picture naming facilitation if the ending syllable prime was given hundreds of milliseconds after picture

onset, supporting a serial word form retrieval during picture naming (Wilshire et al., 2016). Thus, having the design of a covert picture naming task with ending phoneme judgment can maximize the likelihood of a complete word form activation without overt output. Wei et al. (2022) also demonstrated that adding gamification elements (compared to traditional experiments) can promote engagement in online experiments and potentially mitigate the boredom and inattention inherent to remote/unsupervised testing conditions (Brehman et al., 2009; Semmelmann & Weigelt, 2017), with greater engagement in the task behaviorally reflected in an overall faster RT (Wei et al., 2022). Moreover, PWI gamification can be an ideal assessment that measures both naming accuracy and reaction time (RT) of the word retrieval processes. The measurement of both accuracy and RT is potentially more sensitive for word-finding difficulties than the mere focus on word-finding failures in confrontational naming tests included in traditional clinical assessments such as the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) and Mini-Mental State Examination (MMSE) (Folstein et al., 1975).

1.5. Current study

This study tests the transmission deficit, inhibition deficit, and processing speed hypotheses on age-related word-finding through three levels of investigation. First, utilizing the gamified picture-word interference task (Wei et al., 2022), we fitted multiple linear regression models to measure the effects of distractor condition and age on naming performance to reveal how the phonological facilitation, categorical interference, and overall picture naming performance change with age across the adult lifespan. Next, to see if the picture-word interference game can index word-finding abilities in real-life natural speech, we fitted multiple regression models to explain picture naming performance with the effects of distractor condition, age, and speech disfluency obtained from automated analysis of naturalistic speech samples. Speech disfluencies characterized in connected speech, usually induced through picture description or prompted conversations, have been identified as one of the most prominent speech aspects that differentiate older from younger adults (Bortfeld et al., 2001; Castro & James, 2014; Horton et al., 2010), and identify mild cognitive impairment and AD patients (e.g., see reviews Filiou et al., 2020; Mueller et al., 2018; Slegers et al., 2018). Moreover, Yeung et al. (2021) characterized a number of speech features and validated them with clinician ratings of different aspects of connected speech. Amongst all speech aspects, features that point to word-finding difficulty, such as rate of speech, word duration, and length and the number of unfilled (silent) pauses, have the highest correlation with clinician ratings (Yeung et al., 2021). Utilizing these word-finding measures from natural speech, we explored the relationship between participants' picture naming performance in the game and their speech disfluencies measured in picture description recordings. Last but not least, to investigate the information universal prediction that language-domain deficits stem from a more general decline in cognitive control, regression models were also used to assess executive functions' effect on picture naming game performance, on top of the effects of age and motor speed.

Assuming that no tip-of-the-tongue state is being elicited during picture naming, the same phonological primes are used across age groups, and the grammatical category of the primes matches the target, the transmission deficit hypothesis could have two

different predictions for the degree of PF across age. Since phonological priming operates through a bottom-up converging activation (many-to-one), this direction of activation is less vulnerable to age-related failure although all node-to-node connections are weakened by age (Mackay, 1988; MacKay & Burke, 1990). Therefore, older adults may either benefit from phonological primes to the same extent as young adults (Taylor & Burke, 2002), or to a lesser extent since the node connection strength is still weakened by age (MacKay & Burke, 1990). Meanwhile, older adults likely have an enriched semantic network through life experience with more extensive shared features amongst conceptually similar words, supported by evidence on stronger semantic priming effect during word recognition in older adults (see meta-analysis Laver & Burke, 1993). Therefore, the transmission deficit hypothesis would predict a larger degree of CI effect from categorical distractors with older age due to increased priming toward semantically-related words. If CI increases with age and/or PF decreases with age as the transmission deficit hypothesis would predict, the amount of that increase/decrease should be associated with the degree of disfluency in natural speech, thereby supporting the assertion that transmission deficit is the reason behind age-related word-finding struggles.

On the other hand, the inhibition deficit hypothesis would predict that older adults not only experience interference from categorical distractors, but also benefit from phonological primes more than young adults, due to their increased susceptibility to distraction (Healey et al., 2008; Zacks & Hasher, 1994). Moreover, if an inhibition deficit is the reason behind age-related word-finding difficulties, we expect the increase of categorical interference and phonological facilitation to correlate with disfluency in natural speech. Alternatively, if general slowing in processing speed is the cause of age-related word-finding difficulty, we would anticipate the overall picture naming RT to increase with age as well as being associated with disfluencies in natural speech after controlling for non-cognitive slowing such as motor reaction time. Furthermore, the inhibition deficit and processing speed theories, advocating an information-universal mechanism behind different domains of cognitive aging, would predict that the degree of CI (Lustig et al., 2001; May et al., 1999), and overall picture naming RT (T. Salthouse, 1996), will be related to levels of executive functioning.

2. Material and method

2.1. Data availability statement

We report all data exclusion, all manipulation, and all measures in the study. All data, materials, codes, and supplementary materials, behind this analysis have been made publicly available on Open Science Framework (https://osf.io/yc7jm/?view_only=40bbaf6d13314b0dbee81ce41b0aa757). Data were analyzed using R, version 4.1.2 (R Core Team, 2021). This study's design and its analysis were not pre-registered.

2.2. Participants

One hundred and eighty-six North American English native speakers were recruited via Prolific (www.prolific.co), the University of Toronto participant pool, and the Baycrest Hospital participant database. Fifteen participants were excluded due to technical issues (e.g., video

call cutoff, WIFI drop off, experimental software glitches), two participants were excluded due to preexisting neurological disorders, one participant who fell asleep halfway through participation was excluded, and two participants were excluded due to low MoCA scores (<21). Note that we cannot exclude the possibility of including some older adults with mild cognitive impairment or prodromal AD since MoCA is not a foolproof screening test. The Carson et al. (2018) and Waldron-Perrine & Axelrod (2012) noted that a significant number of false positives are commonly observed with the traditional 26/30 MoCA cutoff. Thus the lower MoCA cutoff of 21/30 was used in this study. For confirmation, data analyses with a 26/30 MoCA cutoff score are included in Supplementary 4.1, but they do not contradict the results reported below. To control for the interference level of the audio distractors (in the PWI game) intrinsic to the hardware, all participants without headphones or earphones were excluded. This latter criterion excluded 5 young (age 18–35), 4 middle (age 36–63), and 32 older adults (age 64–90). Many older adults, especially those above 75-years-old, did not use headphones or earphones due to discomfort or incompatibility with their hearing aids. The final sample consisted of 125 participants − 36 young (age $M = 26$; SD = 5.63; sex: 22 females), 38 middle-aged (age $M = 48.34$; SD = 8.61; sex: 24 females), and 51 older (age $M = 70.43$; SD = 5.22; sex: 28 females) adults with normal or corrected to normal vision and hearing, and without neurological or psychological disorders. The percentage of biological females in each age group was roughly matched (young: 61%, middle: 63%, older: 54% female) as indicated by a chi-squared test of the association between sex and age group, which was found to be non-significant (Chi-Squared[2] = 0.69, $p = .71$). The remaining older adults had a mean MoCA score of 26.24 (SD = 2.16). Overall, 10 out of 36 young (28%), 4 out of 38 (11%) middle, and 13 out of 51 (25%) older adults were multilinguals, forming comparable percentages between young and older age groups while the middle-aged group had a lower percentage of multilingualism. Given the multilingualism pattern in our participants, any U-shaped relationship with age should be interpreted with caution. Young ($M = 15.94 \pm 2.67$), middle ($M = 16.55 \pm 2.13$), and older ($M = 16.14 \pm 2.24$) participant groups had matched years of education ($F(2,122) = .67$, $p = .52$). Moreover, semantic knowledge as measured by the categorical fluency task (Delis et al., 2001; Fine & Delis, 2011; Shao et al., 2014; Zemla, 2022) did not differ across age ($r = -0.094$, $R^2 = 0.0088$, $p = .30$) (see Table 2).

A priori power analyses were conducted using G*power version 3.1.9.6 (Faul et al., 2007) for sample size estimation for the three steps of investigations in the current study. First, based on data from T. A. Salthouse and Meinz (1995) ($N = 242$), which revealed significant age-related changes in interference ($r = .47$) and facilitation effects ($r = .36$) during Stroop-like tasks, minimum sample sizes of 53 and 94 were needed to reveal age-related changes in distractor interference and facilitation effects with $\alpha = 0.5$ and power = .80. Second, to the best of our knowledge, the literature investigating the association between picture naming and connected speech fluency in healthy adults is very limited; thus, sample size estimation for this research question is difficult to implement. Nonetheless, this research question has been investigated frequently in post-stroke aphasia. According to Mayer and Murray (2003) ($N = 14$), a number of connected speech features reflecting word finding difficulties were related to picture naming task performance in aphasic patients with the lowest correlation being $r = .68$.With $\alpha = 0.5$, and power = .80, the minimum sample size needed for this effect size is 14. Third, we consulted the study of Higby et al. (2019) ($N = 305$), which examined the relationship between a number of executive function tasks and picture naming performance. The different executive composite scores had significant correlations with the

Boston Naming task RT ranging from .45 to .15, with an average correlation of .31. Using the average $r = .31$, $\alpha = 0.5$, and power = .80, the minimum sample size needed for this effect size is 79. Thus, the final sample size of the current study ($N = 125$) is greater than the samples sizes suggested by effect sizes previously found in related research. The Research Ethics Boards of Baycrest Hospital and the University of Toronto approved the study, and all methods were carried out in accordance with Health Canada guidelines and regulations. Informed consent was obtained from all participants electronically, using an online survey platform (project-redcap.org).

2.3. Procedure and measures

Participants started their participation through a video call on the Zoom videoconferencing software. Once they joined the video call, the experimenter sent participants a google form that laid out all the subcomponents of the study and walked them through it step-by-step. After confirming their eligibility, participants clicked on a RedCap link for the informed consent form. The experimenter explained the study, and the participants gave their verbal consent and checked the corresponding statements on the form. After the informed consent, participants carried out 3 phases of the study sequentially. Phase 1 was the PWI game with covert naming. Phase 2 involved two natural speech sample elicitations through a picture description task. Phase 3 included administration of a set of executive function measures, consisting of a reading span task, a motor and perceptual inhibition task, and verbal fluency tasks. For older adults, their cognitive ability was assessed using the Montreal Cognitive Assessment (MoCA) test (Nasreddine et al., 2005) between phases 1 and 2. The order of the phases/tasks was designed to avoid carry-over effects on word-generation tasks. Since the PWI game, the reading span task, and the verbal fluency tasks required generation and memorization of words, other tasks (picture description, motor and perceptual inhibition tasks) were inserted between them to reduce potential inter-task interference.

2.3.1. Phase 1: Picture-Word Interference (PWI) game

The PWI game included in the current study was adopted from Wei et al. (2022), who observed CI and PF effects during covert picture naming at similar picture-word SOAs as those observed in other psycholinguistic studies employing overt naming with vocal reaction time as the outcome measure (see review in Abel et al., 2009). Specifically, a CI effect was induced using an earlier SOA, when an auditory distractor word precedes the target picture (i.e., −200 SOA), while a PF effect was induced when an auditory distractor word and target picture were presented simultaneously (i.e., 0 SOA). In the PWI game, participants covertly named pictures of common objects. Each trial was composed of a 500 ms fixation cross and a 2500 ms picture presentation in the middle of the screen. Participants were instructed to ignore the audio words they heard and to focus on judging whether the picture names ended in the target sound assigned for the block by "whacking" the "yes" or "no" mole through a single key press (see Figure 1). The speed and accuracy of the button responses were the target measurements. At the beginning of each block, a target phoneme (i.e., /k/,/l/,/n/,/r/,/s/, or/t/) was presented to the participants

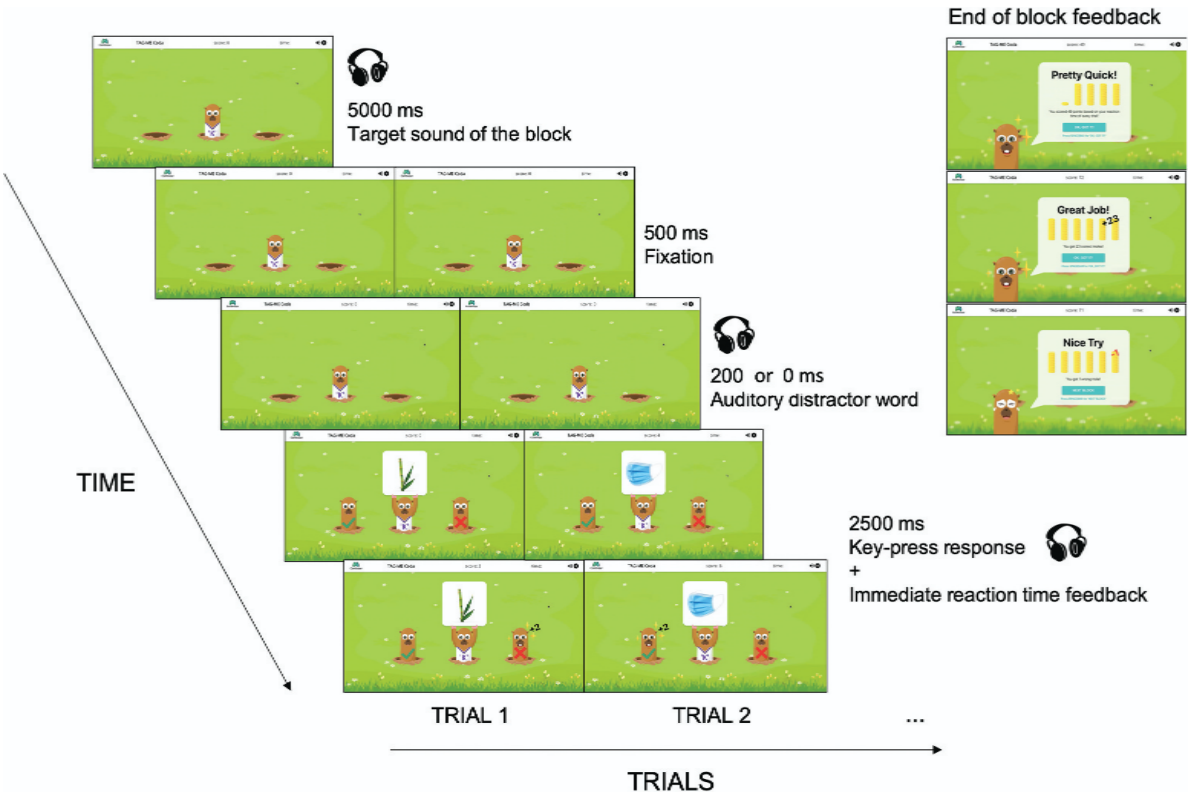


Figure 1. Picture-word interference game.

both visually and auditorily for 5000 ms. The experiment contained 6 blocks in total, and each block had 24 trials, totaling 144 trials. The trials included 4 different picture-word conditions adopted from Wei et al. (2022) – unrelated condition with −200 ms SOA (UN_200), unrelated condition with 0 ms SOA (UN_0), categorically-related condition with −200 ms SOA (C_200), phonologically-related condition with 0 ms SOA (P_0). The two unrelated conditions were set at different SOAs to serve as the respective baseline condition for the categorical and phonological conditions with matching SOA. The picture-word pairs that elicited the strongest categorical slowing (C_200_{RT} – UN_200_{RT}) and phonological speeding (UN_0_{RT} – P_0_{RT}) effects in the Wei et al. (2022) study 4 were included in this study (see Supplementary 1.1 for the picture-word pairs). The 36 trials for each condition were evenly split into 6 blocks, with each block containing 6 trials per condition.

The picture-word trials under each target phoneme block were put into 4 pre-randomized lists and counterbalanced across participants. Each picture was paired with a unique audio distractor word, and was therefore limited to one experimental condition. A fully balanced design (e.g., Latin Squares) was not used because many pictures did not have suitable distractor words available for multiple conditions (i.e., phonological and categorical) that met our criteria of high frequency and low syllable count. Kruskal – Wallis tests indicated that the syllable count (ranging from 1 to 4 syllables) and word frequency of the pictures were matched between the 4 picture-word conditions (**syllable count** $H(3) = .91$, $p = .82$; **word frequency** $H(3) = 4.531$, $p = .21$;). The speech duration ranged between 720 and 1,300 ms (mean 937 ms) and was matched between conditions ($H(3) = 3.5$, $p = 0.32$). Under each block,

half of the trials per condition were "yes" trials (picture name ends in the target phoneme) and the other half were "no" trials (picture name does not end in the target phoneme). The distractor words never had the same ending phoneme as the paired target pictures. The ending phoneme judgment was made by pressing the "yes" and "no" keys (corresponding to the "F" and "J" keys) on the keyboard as quickly as possible within a maximum reaction time of 2500 ms. The mapping of "yes" and "no" keys' to the left (the "F" key) and right (the "J" key) hands, or vice versa, was counterbalanced across participants. Although we cannot control for each participant's hardware refresh rate, it ranged between 60–240 Hz, mostly at multiples of 60 Hz. All the stimuli (pictures and audio files) were pre-downloaded by the website and all participants were required to use Chrome as their web-browser. To control for the auditory volume of the distractors, participants went through a headphone/earphone volume calibration procedure (Woods et al., 2017) to set the volume to a subjectively loud but comfortable level before starting the game.

2.3.2. Phase 2: picture descriptions

Two black-and-white drawings depicting complex scenarios were presented to the participants one by one through the share-screen function on Zoom. These drawings are similar to the "Cookie Theft" picture description task from the widely-used Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983) but original and proprietary to Winterlight Labs, developers of the speech analysis software used in this study, and the exact pictures (See Figure 2) have been validated in multiple studies detecting pathological aging such as AD (Curcic et al., 2022; Robin et al., 2021). Participants were instructed to describe everything they saw in the pictures until we asked them to stop. They described the events in each picture in natural speech for 60 seconds, 2 pictures in total. Analyses of naturalistic speech samples have been used to develop natural language algorithms that extract acoustic and linguistic features, which can index language abilities such as semantic and syntactic complexity, discourse repetition, speech coherency, word-finding difficulties, etc (see Fraser et al., 2015; Robin et al., 2021). The picture description speech samples recorded from each participant were manually transcribed by professionally trained transcribers. Data processing and feature extraction were performed using the Winterlight Labs pipeline (www.winterlightlabs.com) with Python-

Figure 2. Winterlight Labs© picture description stimuli.

based standard acoustic and language processing libraries and proprietary custom code. The number of non-identifiable words, average word duration, number of filled pauses, hesitation, long pause (>2s) count, medium pause (1-2s) count, pause word ratio, short pause (<1s) count, short pause duration, speech rate, number of 'uh's and 'um's, and number of unfilled pauses were extracted from their picture description audio and transcription files (see Supplementary 2.1 for the definition of these features). These 13 speech features have been categorized as indicative of the degree of speech disfluencies/word-finding difficulty (Rohrer et al., 2008; Yeung et al., 2021).

2.3.3. Phase 3: Executive function measures

2.3.3.1. Reading span. The reading span task was adapted from Klaus and Schriefers (2006), created using the PsychoPy Builder interface (Peirce et al., 2019), output to a PsychoJS experiment (Bridges et al., 2020), and implemented online through Pavlovia (https://pavlovia.org/). In each trial of the task, participants saw a sentence to read out loud and judged whether it was sensical (e.g., "People in our town are more giving and cheerful at Christmas time.") or non-sensical (e.g., "During the week of final spaghetti I felt like I was losing my mind.") within 10 seconds. Following the sentence judgment, a word showed up in the middle of the screen for them to read out loud and memorize. Half of the time, the words for memorization were concrete nouns, while the other times they were abstract nouns. Each pair of sentence judgment and word memorization is a trial, and multiple trials consisted of a set. The set size increased incrementally from 2 to 6 trials over the course of the task, and each set size was repeated twice before moving on to the next set size. At the end of each set, participants were prompted to type out all the words that they still remember in the set. This task measures working memory capacity as the ability to remember a rolling list of words while actively processing sentences (Daneman & Carpenter,1980; Gick et al., 1988). The working memory score was calculated using the proportion words method suggested by Friedman and Miyake (2005). For example, remembering 2 out of 5 words in a set gets 0.4 points. The mean score over all sets was the final reading span score.

2.3.3.2. Motor and perceptual inhibition. This task adapted the design by Nassauer and Halperin (2003) to disassociate two types of cognitive inhibition – termed motor inhibition and perceptual inhibition (see also Germain & Collette, 2008; Jennings et al., 2011). This task was created using the Psychopy Builder interface (Peirce et al., 2019), output as a PsychoJS experiment, and implemented online through Pavlovia (https://pavlovia.org/). To control for the dimensions of each participant's screen display, participants adjusted a standard credit card image on the screen to match with a physical credit card in their hands (8.6 cm wide and 5.4 cm tall; Morys-Carter, 2020). These measured parameters allowed Psychopy to scale the stimuli and their location consistently across all participants. Then, participants were instructed to keep their left hand, and right hand, index fingers on the keyboard "F" and "J" keys, respectively. This task was composed of 5 runs. In run 1, an arrow pointing either left or right showed up in the middle of the screen. Participants were asked to press the keys according to the arrow-pointing direction as quickly as possible. In run 2, a rectangle showed up on either the left or right side of the screen. Participants were instructed to press the keys according to the rectangle location as quickly as possible, measuring simple motor RT. In run 3, an arrow pointing either left or

right showed up on either the left or right side of the screen, creating the congruent (i.e., when the arrow-pointing direction is the same as the location) and incongruent (i.e., when the arrow-pointing direction is the opposite of the location) conditions. Participants were instructed to ignore the location of the arrow but respond according to the arrow-pointing direction. Run 4 repeated run 1 to re-familiarize participants to respond to the arrow-pointing direction. Lastly, in run 5, an arrow pointing either left or right showed up in the middle of the screen. However, in contrast to runs 1 and 4, participants needed to press the keys corresponding to the opposite direction of the arrow-pointing direction. Perceptual inhibition (PI) was calculated as the RT difference between the congruent and incongruent conditions in run 3. Perceptual incongruent inhibition (PIi) was calculated as the RT difference between the incongruent condition in run 3 and the no-conflict condition in runs 1 and 4. On the other hand, motor inhibition (MI) was calculated as the RT difference between reacting to the opposite of the arrow-pointing direction (run 5) versus simply reacting to the arrow-pointing direction (runs 1 and 4). PI and PIi measured the ability to focus attention when interfering stimuli are present while MI represented the ability to suppress the incorrect response that may be triggered by the presented stimulus (Germain & Collette, 2008; Jennings et al., 2011; Nassauer & Halperin, 2003).

2.3.3.3. Verbal fluency. This task was based on the DKEFS battery (Delis et al., 2001), implemented through verbal instructions given in the video call. The DKEFS battery has three conditions – letter fluency, categorical fluency, and categorical switching. In each condition, two prompts were given to the participants. In the letter fluency condition, participants were asked to generate as many words in 60 seconds as possible based on a starting letter provided to them. This task was conducted first with the letter "F" and then with "A". In the categorical fluency condition, participants were asked to think of as many animals' and boys' names as possible within 60 seconds each. In the categorical switching condition, participants needed to alternate between fruits and pieces of furniture (e.g., apple, table, pear, chair . . .), naming as many members of the two categories as possible within 60 seconds. This task measured participants' ability to generate words fluently in an effortful, phonemic format (letter fluency), from overlearned concepts (categorical fluency), and shifting between overlearned concepts (categorical switching) while self-monitoring for the rules and restrictions. The number of times participants deviated from the rules and the number of repetitions was also recorded as error measures. Letter fluency can reveal participants' verbal knowledge, systematic retrieval of lexical items, and simultaneous processing and monitoring (Delis et al., 2001; Fine & Delis, 2011; Shao et al., 2014). Categorical fluency can measure participants' semantic network/knowledge and categorical switches also index shifting ability (Delis et al., 2001; Fine & Delis, 2011; Shao et al., 2014; Zemla, 2022).

3. Results

Descriptive data on the game and executive measures and their correlations with age are summarized in Tables 1 and 2, respectively.

For the multiple regression models fitted for the 3 steps of the investigation (i.e., game performance explained by age, by natural speech disfluency, and by executive function), to control for type I error during multiple comparison, we corrected our alpha level using

Table 1. Descriptive statistics for game measures, showing means and standard deviations in parentheses, as well as correlations with age.

Game measures	Young adults (N = 36)	Middle-aged adults (N = 38)	Older adults (N = 52)	Correlation with age (N = 126)
Accuracy-based (%)				
C_200_{ACC}	85.95 (7.15)	79.58 (12.98)	72.73 (13.65)	r = −.50, R^2 = .25, $p < 0.001$***
P_0_{ACC}	94.91 (4.58)	91.81 (8.99)	86.99 (12.15)	r = −.39, R^2 = .15, $p < 0.001$***
UN_200_{ACC}	94.06 (4.89)	89.16 (10.29)	85.03 (10.82)	r = −.42, R^2 = .18, $p < 0.001$***
UN_0_{ACC}	88.66 (7.29)	84.82 (14.53)	80.7 (11.91)	r = −.35, R^2 = .12, $p < 0.001$***
$AllCond_{ACC}$	90.89 (4.56)	86.34 (10.68)	81.36 (10.92)	r = −.46, R^2 = .21, $p < 0.001$***
RT-based (ms)				
C_200_{RT}	1331.8 (184.53)	1435.51 (268.62)	1650.28 (226.43)	r = .54, R^2 = .29, $p < 0.001$***
P_0_{RT}	1131.96 (167.3)	1259.49 (266.03)	1490.91 (243.15)	r = .58, R^2 = .34, $p < 0.001$***
UN_200_{RT}	1132.11 (166.07)	1237.87 (237.09)	1476.49 (235.77)	r = .59, R^2 = .35, $p < 0.001$***
UN_0_{RT}	1339.05 (191.67)	1428.63 (261.74)	1618.08 (241.56)	r = .50, R^2 = .25, $p < 0.001$***
$AllCond_{RT}$	1233.73 (168.92)	1340.37 (250.84)	1558.94 (227.87)	r = .57, R^2 = .32, $p < 0.001$***

Notes. $AllCond_{ACC}$ and $AllCond_{RT}$ are the mean accuracy and RT, taken as the overall picture-naming accuracy and RT, across all distractor conditions (i.e., C_200, P_0, UN_200, UN_0). Pearson's correlation coefficient was used for correlations between executive measures and age.

Table 2. Descriptive statistics for executive measures, showing means and standard deviations in parentheses, as well as correlations with age.

Executive measures	Young adults (N = 36)	Middle-aged adults (N = 38)	Older adults (N = 52)	Correlation with age (N = 126)
Reading span (working memory)				
Proportional word recall	.70 (.11)	.64 (.14)	.58 (.14)	r = −.46, R^2 = .21, $p < 0.001$***
Sentence accuracy (%)	93.61(3.57)	94.52 (2.82)	93.33 (4.01)	r = −0.064, R^2 = 0.0041, p = .48
Sentence RT (s)	4.68 (.70)	4.97 (.84)	4.89 (.65)	r = .12, R^2 = 0.015, p = .17
Motor and perceptual inhibition (inhibition control)				
Simple detection RT (ms)	367.43 (71.82)	398.69 (67.50)	458.5 (91.18)	r = .48, R^2 = .23, $p < 0.001$***
MI (ms)	82.64 (82.36)	144.6 (121.07)	172.29 (149.92)	r = .32, R^2 = .10, $p < 0.001$***
PI (ms)	30.8 (47.40)	52.01 (45.77)	75.35 (72.09)	r = .32, R^2 = .10, $p < 0.001$***
PIi (ms)	154.79 (71.73)	190.78 (96.41)	235.78 (127.35)	r = .32, R^2 = .10, $p < 0.001$***
Verbal fluency (vocabulary, switching, self-monitoring)				
Letter fluency	47.03 (13.07)	49.03 (14.13)	49.22 (13.00)	r = 0.061, R^2 = 0.0037, p = .50
Categorical fluency	46.75 (8.27)	46.71 (12.44)	44.82 (10.83)	r = −0.094, R^2 = 0.0088, p = .30
Categorical switching total correct	16.03 (3.18)	16.26 (2.86)	14.63 (2.89)	r = −.19, R^2 = 0.036, $p < 0.05$*
Categorical switching total switch	15.58 (3.27)	15.42 (2.98)	14.2 (2.95)	r = −0.18, R^2 = 0.032, $p < 0.05$*
Total set loss	.78 (1.15)	1.18 (1.74)	1.61 (1.43)	r = .21, R^2 = 0.045, $p < 0.05$*
Total repetition	1 (.96)	1.03 (1.17)	2.92 (2.82)	r = .25, R^2 = 0.061, $p < 0.01$**

Notes. Pearson's correlation coefficient was used to calculate correlations reported in this table.

Bonferroni correction (0.05/number of tests). Since the CI and PF effects should be elicited through different game conditions, separate models were implemented with the 0 SOA conditions (P_0, UN_0) and the −200 SOA conditions (C_200, UN_200) as predictors for game performance (RT or accuracy). Therefore, each investigation involved fitting 4 separate multiple regression models, 2 game measures (RT & accuracy) x 2 game condition SOAs (0 and −200), the corrected p-value threshold was set at 0.0125.

3.1. Game performance explained by age

To investigate the categorical interference (CI) and phonological facilitation (PF) effects, as well as their changes with age, a series of multiple linear regression models were conducted to predict picture naming game performance with age, game condition, and their interaction while controlling for simple motor RT. The models were fitted in R using the aov() function (R Core Team, 2021) as shown below, with the error term specifying game condition as a within-subject variable:

0 SOA : *aov(game measure ~ age * game condition (UN_0 VS. P_0) + simple motor RT + Error (subject/game condition))*
-200 SOA : *aov(game measure ~ age * game condition (UN_200 vs. C_200) + simple motor RT + Error(subject/game condition))*

The simple motor RT was measured in session 2 of the motor and perceptual inhibition task, which required participants to press the corresponding keys by detecting the location of a rectangle as quickly as possible.

As expected, picture naming speed and accuracy decreased with age in all game conditions (**RT 0 SOA**: $F(1) = 61.09$, $p < 0.001$, $\eta^2_p = .33$; **RT −200 SOA**: $F(1) = 70.01$, $p < 0.001$, $\eta^2_p = .36$; **Accuracy 0 SOA**: $F(1) = 21.01$, $p < 0.001$, $\eta^2_p = .15$; **Accuracy −200 SOA**: $F(1) = 40.31$, $p < 0.001$, $\eta^2_p = .25$) (see Figure 3). Moreover, categorical distractors at −200 SOA significantly slowed picture naming speed ($F(1) = 451.42$, $p < 0.001$, $\eta^2_p = .79$) and interfered with picture naming accuracy ($F(1) = 157.35$, $p < 0.001$, $\eta^2_p = .56$) compared to unrelated distractors at the same SOA see Figure 3(a,c). Meanwhile, phonological distractors at 0 SOA significantly sped up picture naming ($F(1) = 359.88$, $p < 0.001$, $\eta^2_p = .75$) and facilitated picture naming accuracy ($F(1) = 81.64$, $p < 0.001$, $\eta^2_p = .40$) compared to unrelated distractors at the same SOA see Figure 3(b,d). See Supplementary Table S4 and 5 for the full model outputs. These results suggested that the PWI game successfully elicited the CI and PF effects, as seen in our previous studies restricted to healthy young adults (Wei et al., 2022). Additionally, significant interactions between game conditions and age in predicting picture naming speed and accuracy indicated that the magnitude of CI and PF effects changed across the adult lifespan. The degree of phonological speeding significantly decreased as age increased ($F(1) = 13.11$, $p < 0.001$, $\eta^2_p = .10$), suggesting that older adults benefited less from phonological primes compared to young adults (observe that the distance between the two lines decreases with age in Figure 3(b).

On the other hand, the degree of categorical interference in picture naming accuracy significantly increased with age ($F(1) = 8.43$, $p = 0.004$, $\eta^2_p = 0.06$), showing an increased susceptibility to categorical distractors in older compared to young adults (see the increase of line distance with age in Figure 3(c). In other words, older adults made more naming errors than young adults when categorical distractors were present. Since each trial has a maximum response time of 2500 ms and missed responses were encoded as incorrect, both RT and accuracy are important measures of the picture naming game. Younger adults (age 18–64) had a 2.3% no response rate and older adults (age 64–90) had a 8.4% no response rate. A composite score combining RT and accuracy was computed as an inverse efficiency score (J. Townsend & Ashby, 1983; J. T. Townsend & Ashby, 1978), which also suggested a significant increase of CI with age but an age-invariant PF effect (see Supplementary 2.2). Overall, the results showed that CI increased with age, while PF

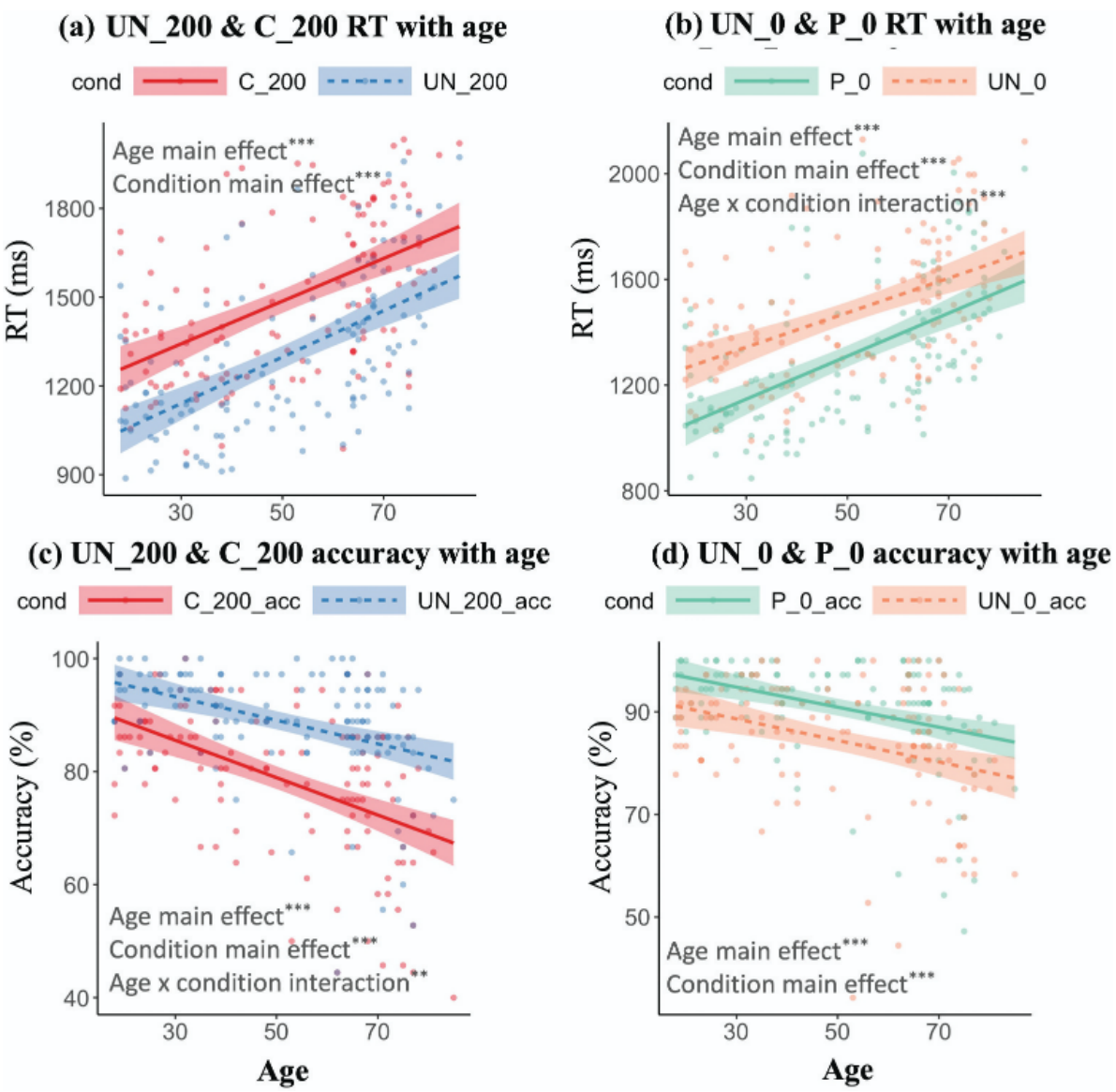


Figure 3. Game RT and accuracy with age by game condition.

either decreased with age when looking at naming RT alone, or remained stable across age when looking at the composite score of RT and accuracy.

3.2. Game performance explained by disfluency measured in natural speech

3.2.1. Factor analysis on speech disfluency features

To decrease feature dimensionality, an exploratory factor analysis using the maximum likelihood factor solution was implemented on the 13 speech disfluency features, plus speech feature scores on long, medium, and short pause durations (a total of 16 features). (See supplementary 2.3.1 for data suitability checks.) Originally, a parallel analysis (Horn, 1965), which generates random correlation matrices and compares their eigenvalues with the eigenvalue of the observed data, suggested 3 factors (see Supplementary 2.4 for the 3-factor model pattern matrix). Eventually, a 2-factor model was chosen for better interpretability. The only difference between the 2-factor and 3-factor models was that the 3-factor model separated longer-pause and shorter-pause-related speech features, while

the 2-factor model simply extracted one pause-related factor. An oblique rotation (oblimin) was performed to allow for inter-factor correlation and the 2 factors were positively correlated (Pearson's $r = .48$, $p < 0.001$), suggesting that more pauses was associated with slower speech speed. The obtained pattern matrix is shown in Table 3.

The first factor had an initial eigenvalue of 6.54 (sum of squared loading of 5.10), and it accounted for 40.89% of the total variance in the data. Factor two had an eigenvalue of 2.49 (sum of squared loading of 2.84) and accounted for a further 15.55% of the total variance. In sum, the two factors accounted for 56.44% of the total variance. Loadings were high on Factor 1 for the eight pausing-related speech features. The higher the factor 1 value, the more pausing tended to occur between speech elements. On the other hand, factor 2 was heavily loaded by the main measures of speech speed. The higher the factor 2 value, the slower the speech rate. In summary, the two factors can be interpreted as underlying aspects of speech disfluency.

3.2.2. Game performance explained by speech disfluency factors

With the underlying factors for speech disfluency extracted, the associations between these factors and game performance were investigated using a series of multiple linear regression models. Each participant's factor score on the 2 disfluency factors were estimated by multiplying the data matrix with the weighting matrix calculated with the default "Thurstone" regression method using the fa() function of the "psych" package (Revelle, 2021) in R. In other words, the factor scores represent how high or low a participant scores (i.e., relative spacing/standing) on the latent factors. In each model, a game measure (RT or accuracy) was specified as the predicted value, while the predictors were the disfluency factor score, game condition, age, simple motor RT, and the 2-way and 3-way interactions between disfluency factor, game condition, and age. The focuses of the results were the main effect of disfluency factor, and its interactions with game condition and age on predicting naming game performance, while controlling for the main effect of simple motor RT. These analyses can reveal whether the disfluency factors were associated with picture naming game RT or accuracy, and whether the associations differ across age or

Table 3. Two-factor model on disfluency speech features.

	Factors		
Features	1	2	Interpretation
Medium pause duration (s)	**1.00**	−0.01	Pausing
Mean pause duration (s)	**.89**	0.08	
Long pause count	**.87**	−0.06	
Long pause duration (s)	**.84**	−0.04	
Medium pause count	**.72**	−0.03	
Short pause duration (s)	**.66**	0.08	
Pause-word ratio	**.62**	.20	
Short pause count	**.36**	−0.01	
Speech rate	0.02	**−1.00**	Speech speed
Average word duration (s)	0.02	**.98**	
Unfilled pause count	.30	**.56**	
Filled pause count	−.33	**.42**	
Hesitation	.20	**.40**	
Non-identifiable word count	.22	−.22	
"uh" count	−.18	.27	
"um" count	−.25	.25	

Notes. Maximum likelihood extraction; Oblimin rotation; Loadings larger than .34 are in bold.

game condition due to CI and PF effects. Again, the 0 SOA and −200 SOA game conditions were specified in separate models. The models were implemented in R as below, with the error term specifying game condition as a within-subject variable:

0 SOA : *aov(game measure ~ speech disfluency factor * game condition (UN_0 VS. P_0) * age + simple motor RT + Error(subject/game condition))*
-200 SOA : *aov(game measure ~ speech disfluency factor * game condition (UN_200 vs. C_200) * age + simple motor RT + Error(subject/game condition))*

The model results suggested that disfluency factor 1 (pausing in speech) did not explain a significant amount of variance in picture naming speed or accuracy with both 0 and −200 SOA game conditions (**RT 0 SOA**: $F(1) = .91$, $p = .34$, $\eta^2_p = 0.01$; **RT −200 SOA**: $F(1) = 1.75$, $p = .19$, $\eta^2_p = 0.01$; **Accuracy 0 SOA**: $F(1) = 0.08$, $p = .78$, $\eta^2_p = 0.00$; **Accuracy −200 SOA**: $F(1) = 0.09$, $p = .76$, $\eta^2_p = 0.00$) (see Figure 4). Neither did disfluency factor 1 (pausing in speech) interact with age (**RT 0 SOA**: $F(1) = .28$, $p = .60$, $\eta^2_p = 0.00$; **RT −200 SOA**: $F(1) = 0.03$, $p = .87$, $\eta^2_p = 0.00$; **Accuracy 0 SOA**: $F(1) = .29$, $p = .59$, $\eta^2_p = 0.00$; **Accuracy −200 SOA**: $F(1) = .42$, $p = .52$, $\eta^2_p = 0.00$) or interact with game condition (**RT 0 SOA**: $F(1) = 0.02$, $p = .88$, $\eta^2_p = 0.00$; **RT −200 SOA**: $F(1) = .74$, $p = .39$, $\eta^2_p = 0.01$; **Accuracy 0 SOA**: $F(1) = 0.04$, $p = .84$, $\eta^2_p = 0.00$; **Accuracy −200 SOA**: $F(1) = 1.00$, $p = .32$, $\eta^2_p = 0.01$) in predicting picture naming speed and accuracy (see Figure 4). See Supplementary Table S6 and 7 for full model outputs with disfluency factor 1 as a predictor.

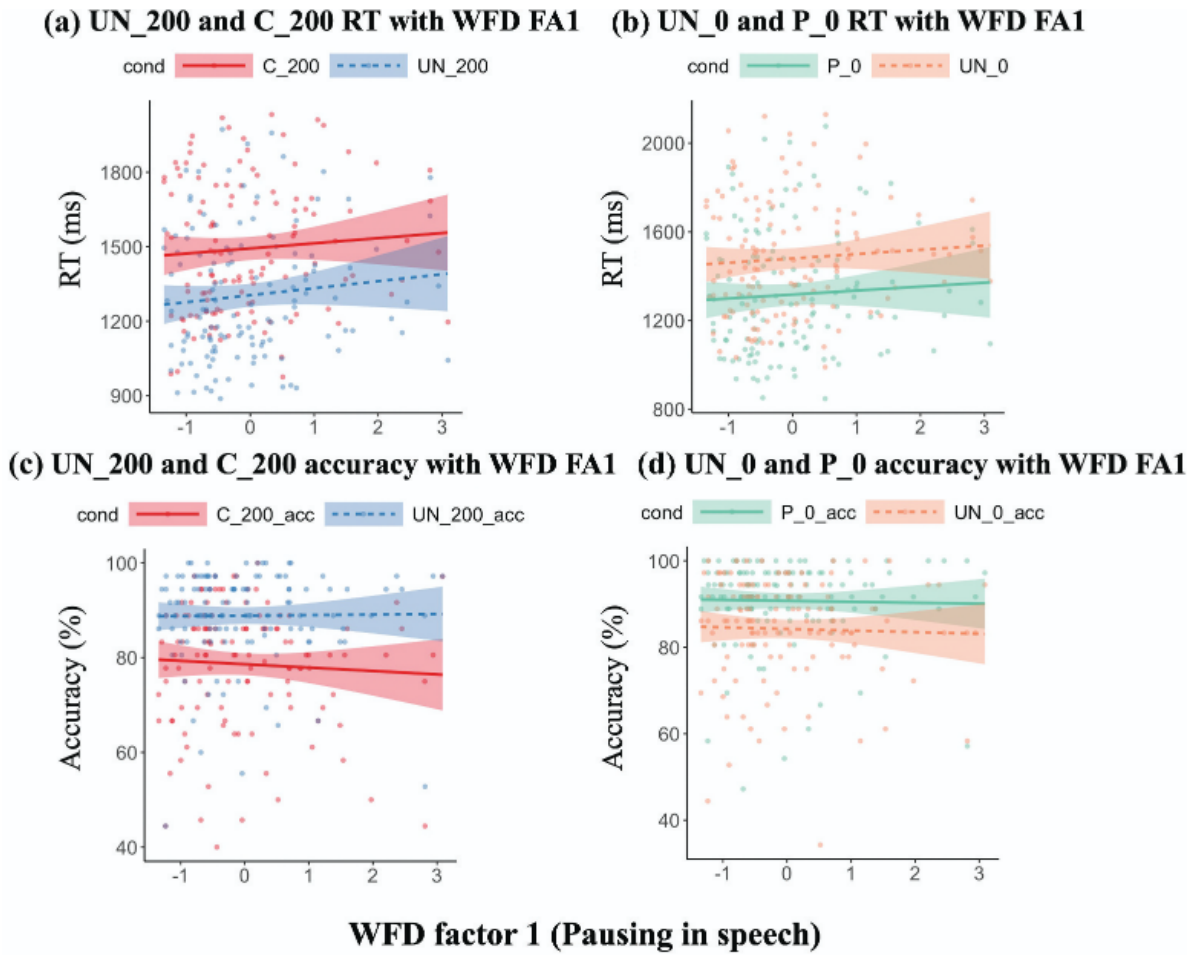


Figure 4. Game RT and accuracy with speech disfluency factor 1 by game condition.

On the other hand, although disfluency factor 2 (speed of speech) did not interact with age (**RT 0 SOA**: $F(1) = .73$, $p = .40$, $\eta^2_p = 0.01$; **RT −200 SOA**: $F(1) = .75$, $p = .39$, $\eta^2_p = 0.01$; **Accuracy 0 SOA**: $F(1) = 0.05$, $p = .82$, $\eta^2_p = 0.00$; **Accuracy −200 SOA**: $F(1) = .13$, $p = .71$, $\eta^2_p = 0.00$) or game condition (**RT 0 SOA**: $F(1) = 0.07$, $p = .79$, $\eta^2_p = 0.00$; **RT −200 SOA**: $F(1) = .12$, $p = .73$, $\eta^2_p = 0.00$; **Accuracy 0 SOA**: $F(1) = .71$, $p = .40$, $\eta^2_p = 0.01$; **Accuracy −200 SOA**: $F(1) = 1.15$, $p = .28$, $\eta^2_p = 0.01$) in explaining picture naming speed or accuracy, it did explain a significant amount of variance in picture naming speed in all game conditions via a main effect of speech disfluency (**RT 0 SOA**: $F(1) = 7.85$, $p < 0.01$, $\eta^2_p = 0.06$; **RT −200 SOA**: $F(1) = 13.06$, $p < 0.001$, $\eta^2_p = .10$) (see Figure 5(a,b), suggesting that slower picture naming RT in the game is associated with slower speech rate in natural speech. Unlike picture naming RT, picture naming accuracy did not associate with speech disfluency factor 2 (speed of speech) (**Accuracy 0 SOA**: $F(1) = 1.85$, $p = .17$, $\eta^2_p = 0.02$; **Accuracy −200 SOA**: $F(1) = 2.81$, $p = .10$, $\eta^2_p = 0.02$) see Figure 5(c,d). See Supplementary Table S8 and 9 for full model outputs with disfluency factor 2 as a predictor. These results suggested that the game condition effects for CI and PF did not interact with either of the speech disfluency factors. In other words, no association between the degree of CI/PF and natural speech disfluency was found. Nonetheless, the speed of speech aspect of disfluency in natural speech can predict picture naming RT, but not naming accuracy, in the game. The lack of association between disfluency in natural speech and picture naming accuracy could potentially result from a ceiling effect

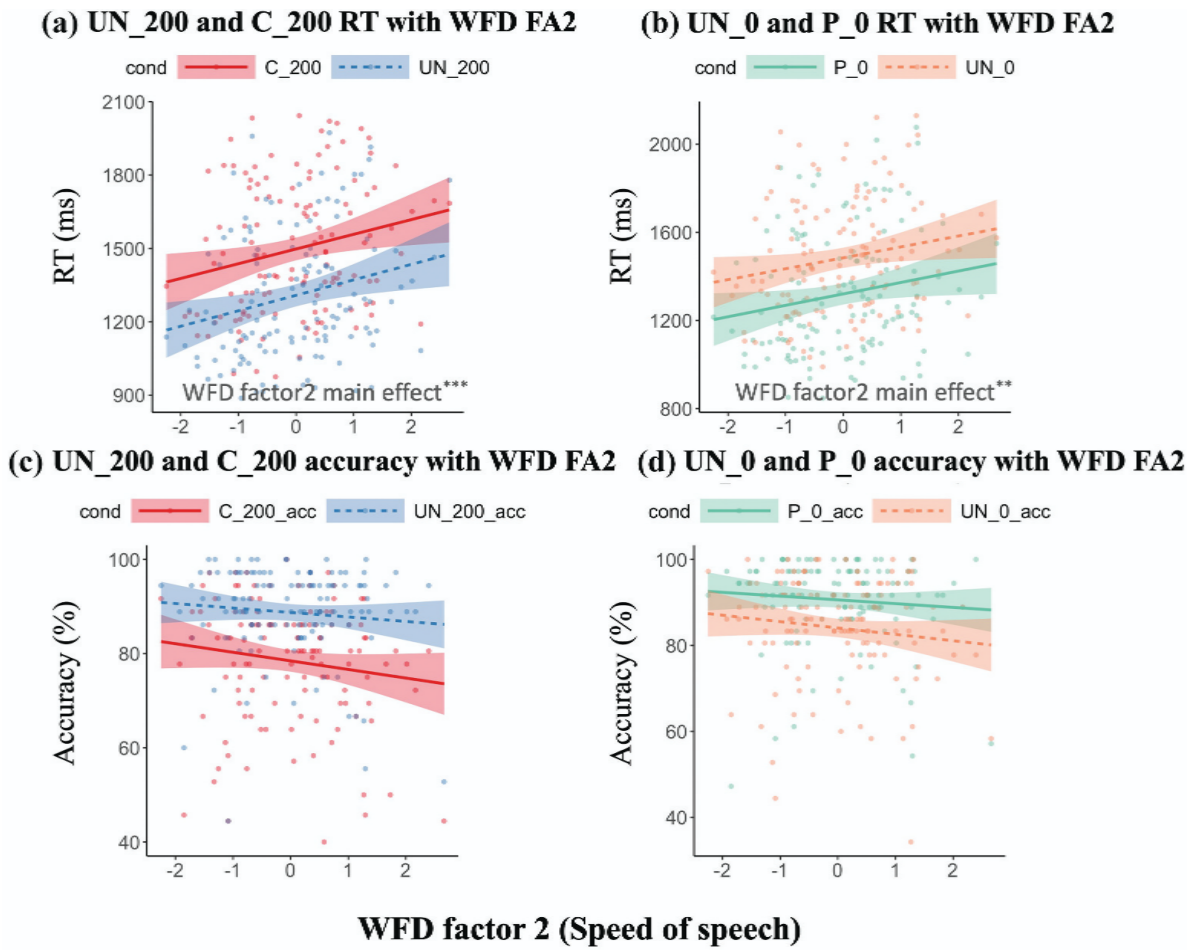


Figure 5. Game RT and accuracy with speech disfluency factor 2 by game condition.

where the majority of participants had accuracy above 85%, as can be expected in a cognitively healthy population. Overall, picture naming RT demonstrated stronger association with natural speech disfluency compared to naming accuracy, CI, and PF measures.

3.3. Game performance explained by executive function

3.3.1. Factor analysis on executive measures

Given that the executive function tests produced a number of output measures, we conducted an exploratory factor analysis using a principal axes factor solution, to see if those measures could be combined into meaningful composites reflecting underlying cognitive abilities. (See supplementary 2.3.2 for data suitability checks.) Initially, a parallel analysis (Horn, 1965) suggested the presence of 4 factors. However, retention of 3 or 4 factors in the model rendered an ultra-Heywood case where the common factors explained more than 100% of the variance in an observed variable, potentially indicating too many common factors were retained and that the factor results were unreliable. Thus, 2 factors were extracted in the exploratory factor analysis in order to maximize interpretability of the results. An oblique rotation (oblimin) was performed to allow for inter-factor correlations and the 2 factors were negatively correlated (Pearson's $r = -.26$, $p < 0.01$), suggested that better verbal fluency, switching, and working memory was related to better inhibition function. The obtained pattern matrix is shown in Table 4.

The first factor had an eigenvalue of 3.23 (sum of squared loading of 2.60), and it accounted for 32.34% of the variance in the data. Factor two had an eigenvalue of 1.78 (sum of squared loading of 1.50) and accounted for a further 17.77% of the variance. In sum, the two factors accounted for 50.11% of the total variance. Verbal fluency, switching, and working memory (WM) related variables had high loadings on Factor 1. Thus, the higher the factor 1 value, the better the verbal fluency, switching, and WM function. On the other hand, factor 2 was associated with inhibition-related variables. Since higher inhibition RT indicates worse inhibitory functioning, high values on Factor 2 indicate lower inhibition ability. Overall, only two items were not associated with the two extracted factors. These items involved failures to self-monitor for instructions, and repetitions.

Table 4. Two-factor model on executive measures.

	Factors		
Features	1	2	Interpretation
Category switching total correct	**.88**	0.03	Verbal fluency, switching, WM
Categorical switching total switch	**.87**	−0.02	
Categorical fluency	**.73**	0.00	
Letter fluency	**.60**	0.00	
Working memory (proportion word recall)	**.30**	−.29	
Perceptual incongruent inhibition (ms)	−0.02	**.81**	Inhibition
Perceptual inhibition (ms)	0.04	**.77**	
Motor inhibition (ms)	−.11	**.35**	
Set loss total count	−.17	−.21	
Repetition total count	.16	−0.02	

Notes. Principal factor extraction; Varimax rotation; Loadings larger than .30 are in bold.

3.3.2. Executive factor effects on game performance

A series of multiple linear regression models were implemented to investigate the effects of executive factor, game condition, age, and their interactions on game performance, while controlling for the effect of simple motor RT. Each participant's factor score on the 2 executive factors were estimated by multiplying the data matrix with the weighting matrix calculated with the default "Thurstone" regression method using fa() function under "psych" package (Revelle, 2021) in R. The predicted value of the model was the specific game measure (RT or accuracy), with the predictors including the specific executive factor score, game condition, age, simple motor RT, and the 2-way and 3-way interactions between executive factor, game condition, and age. These models aimed to reveal associations between executive function and the picture naming game performance, interactions between age and executive function on game performance, as well as associations between executive function and the CI/PF effects (measured through the executive factor x game condition interaction) on game performance. The models were specified in R as below, with the error term specifying game condition as a within-subject variable:

0 SOA : *aov(game measure ~ executive factor * game condition (UN_0 VS. P_0) * age + simple motor RT + Error(subject/game condition))*
-200 SOA : *aov(game measure ~ executive factor * game condition (UN_200 vs. C_200) * age + simple motor RT + Error(subject/game condition))*

When predicting game performance with the first executive factor (verbal fluency, switching, & WM factor), significant effects were observed on picture naming RT and accuracy for all game conditions (**RT 0 SOA**: $F(1) = 67.64$, $p < 0.001$, $\eta^2_p = .36$; **RT −200 SOA**: $F(1) = 67.54$, $p < 0.001$, $\eta^2_p = .36$; **Accuracy 0 SOA**: $F(1) = 23.37$, $p < 0.001$, $\eta^2_p = .16$; **Accuracy −200 SOA**: $F(1) = 30.44$, $p < 0.001$, $\eta^2_p = .20$), showing that better verbal fluency, switching &WM ability is associated with faster and more accurate picture naming (see Figure 6). Furthermore, a significant interaction between the verbal fluency, switching & WM factor and −200 SOA game conditions was found on picture naming accuracy ($F(1) = 11.95$, $p < 0.001$, $\eta^2_p = 0.09$), suggesting that better verbal executive function (fluency, switching, WM) corresponds to weaker categorical interference during picture naming (observe that the distance between the two lines decreases as executive function increases in Figure 6(c). No other significant interaction between game condition and the verbal fluency, switching & WM factor was found (**RT 0 SOA**: $F(1) = 93$, $p = .34$, $\eta^2_p = 0.01$; **RT −200 SOA**: $F(1) = 1.93$, $p = .17$, $\eta^2_p = 0.02$; **Accuracy 0 SOA**: $F(1) = 5.31$, $p = 0.023$, $\eta^2_p = 0.04$) see Figure 6(a-c). No interaction between age and the verbal fluency, switching & WM factor was found in explaining naming performance (**RT 0 SOA**: $F(1) = .11$, $p = .74$, $\eta^2_p = 0.00$; **RT −200 SOA**: $F(1) = .24$, $p = .63$, $\eta^2_p = 0.00$; **Accuracy 0 SOA**: $F(1) = 2.14$, $p = .15$, $\eta^2_p = 0.02$; **Accuracy −200 SOA**: $F(1) = 2.01$, $p = .16$, $\eta^2_p = 0.02$), suggesting independent effects of age and executive function on picture naming. See Supplementary Table S10 and 11 for the full model outputs with executive factor 1 as a predictor.

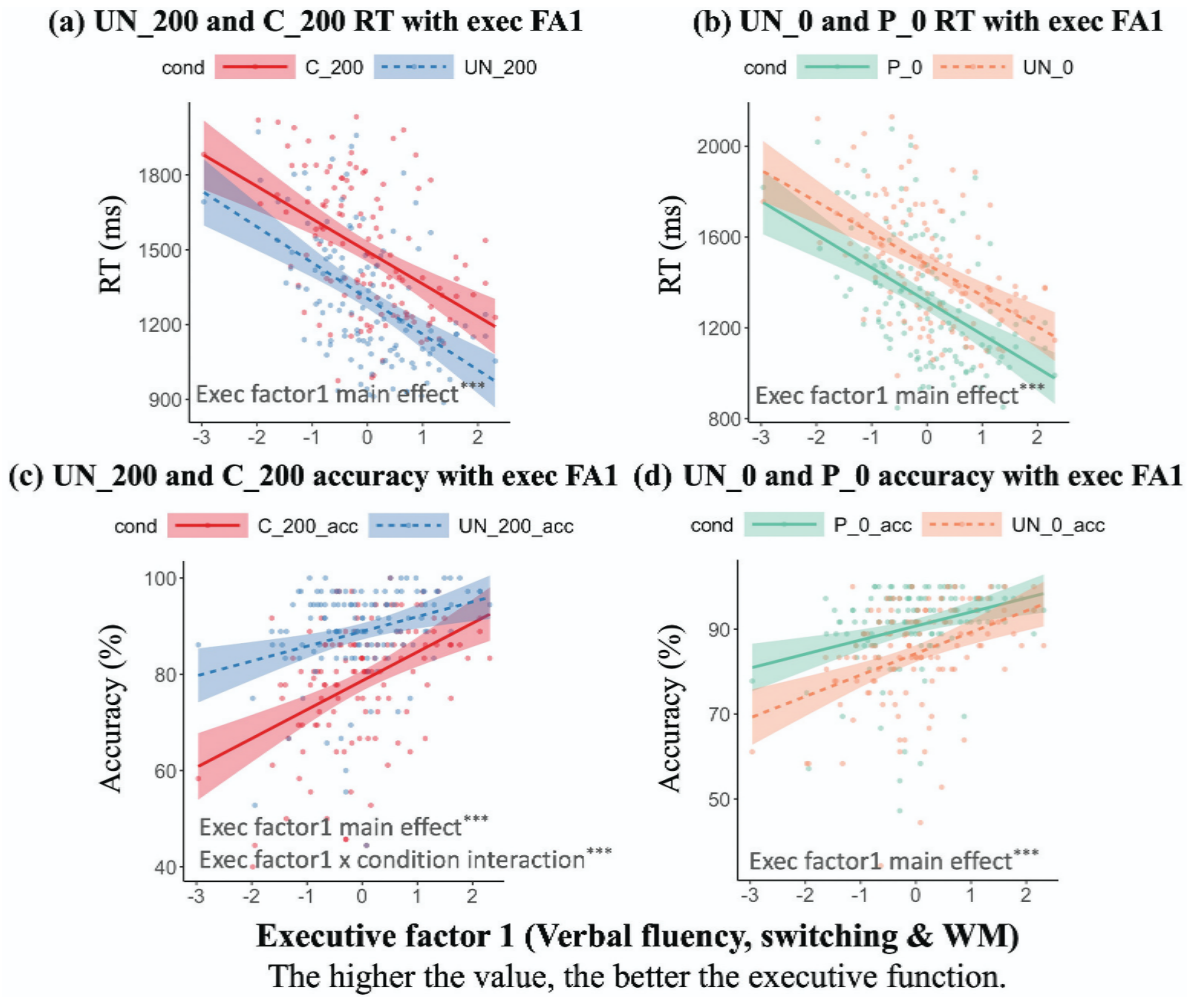


Figure 6. Game RT and accuracy with executive factor 1.

When predicting game performance with the second executive factor (inhibition), significant effects were also observed on picture naming RT and accuracy for all game conditions (**RT 0 SOA**: $F(1) = 55.88$, $p < 0.001$, $\eta^2_p = .32$; **RT −200 SOA**: $F(1) = 42.80$, $p < 0.001$, $\eta^2_p = .26$; **Accuracy 0 SOA**: $F(1) = 14.87$, $p < 0.001$, $\eta^2_p = .11$; **Accuracy −200 SOA**: $F(1) = 13.17$, $p < 0.001$, $\eta^2_p = .10$), showing that better inhibition function is associated with faster and more accurate picture naming (see Figure 7). Inhibition factor did not interact with age (**RT 0 SOA**: $F(1) = 0.00$, $p < 1.00$, $\eta^2_p = 0.00$; **RT −200 SOA**: $F(1) = 0.05$, $p = .82$, $\eta^2_p = 0.00$; **Accuracy 0 SOA**: $F(1) = .91$, $p = .34$, $\eta^2_p = 0.01$; **Accuracy −200 SOA**: $F(1) = .24$, $p = .63$, $\eta^2_p = 0.00$) in explaining game performance, suggesting additive effects of age and inhibition function on picture naming. On the other hand, the inhibition factor had no interaction with game condition in all but one model (**RT −200 SOA**: $F(1) = 4.28$, $p = 0.04$, $\eta^2_p = 0.03$; **Accuracy 0 SOA**: $F(1) = 1.08$, $p = .30$, $\eta^2_p = 0.01$; **Accuracy −200 SOA**: $F(1) = 1.88$, $p = .17$, $\eta^2_p = 0.02$) (see Supplementary Table S12 and 13 for full model outputs). The inhibition factor interacted with the 0 SOA game conditions in predicting picture naming RT in the game ($F(1) = 7.90$, $p < 0.01$, $\eta^2_p = 0.056$) (observe that the two lines converge as the executive factor score increases in Figure 7(b). This interaction suggests that worse inhibition ability (higher factor score) was associated with smaller phonological facilitation during picture naming.

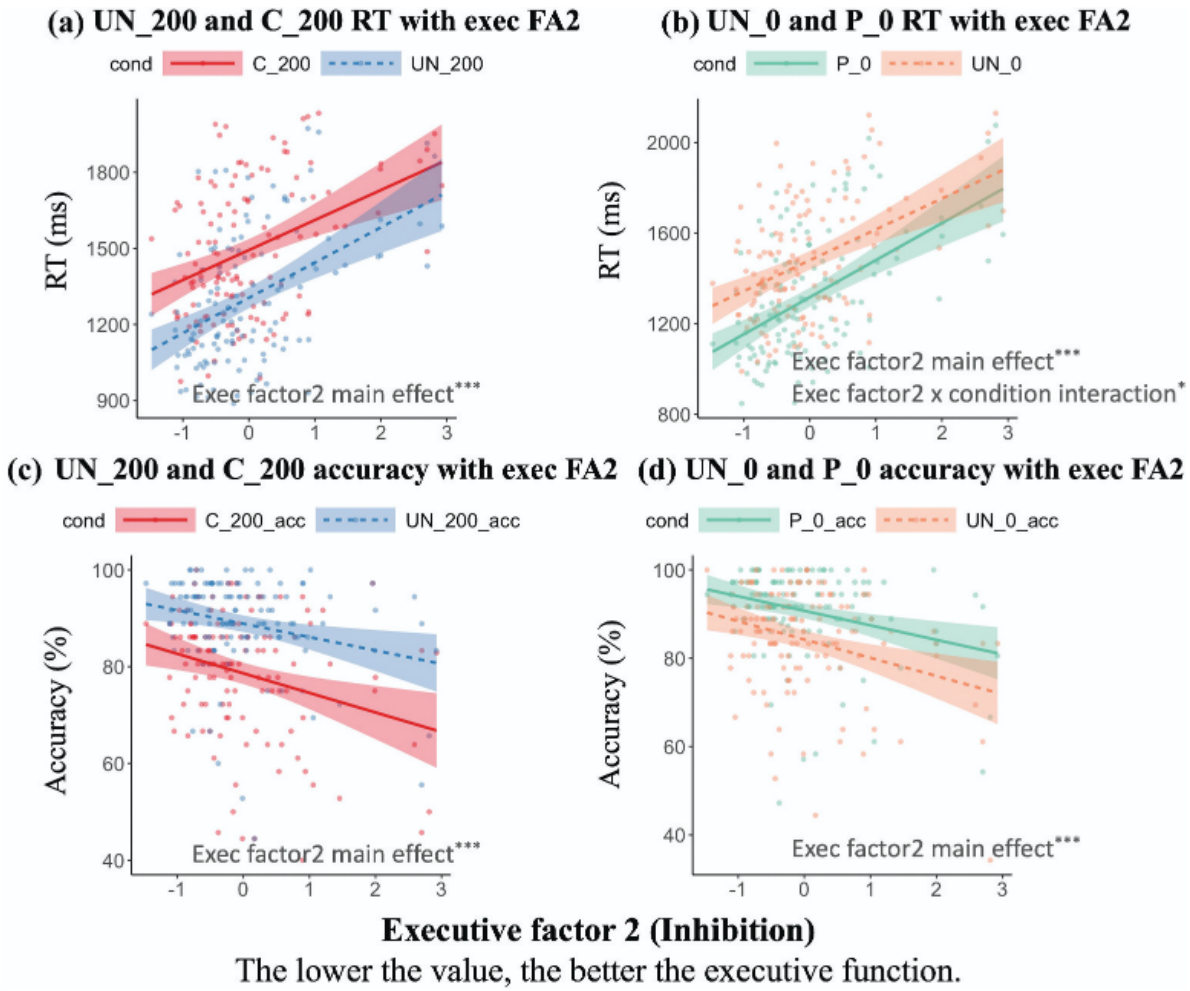


Figure 7. Game RT and accuracy with executive factor 2

4. Discussion

Through the three levels of investigations, this study demonstrated how age, speech disfluency, and executive function interact with the performance of a covert PWI game. The CI and PF effects were successfully induced using the gamification design of the covert PWI task. CI was increased with age, while PF in RT decreased with age and remained invariant across age in the speed and accuracy combined measure. Two speech disfluency factors were extracted from the picture description speech samples. The speed of speech factor significantly explained picture naming RT but not accuracy. Meanwhile, all executive abilities decreased with age. Two executive factors were extracted from the set of executive measures. Both executive factors significantly explained picture naming performance and some CI/PF effects. Further discussion of these findings are framed within their theoretical significance below.

4.1. Age-related changes in PWI supported transmission deficit and processing speed theories

Age-related deficits in phonological activation (MacKay & Burke, 1990), competitor inhibition (Hasher & Zacks, 1988; Zacks & Hasher, 1994), and processing speed (T.

Salthouse, 1996) have been proposed as distinct cognitive mechanisms to account for age-related increases in word-finding difficulty. Three stages of investigation were implemented in the current study to compare these cognitive aging theories. First, using a novel gamified PWI paradigm (Wei et al., 2022), this study demonstrated how measuring the levels of interference from categorical distractors (i.e., CI) and facilitation from phonological primes (i.e., PF) across the adult lifespan allows us to probe the two main stages of word retrieval that may degrade in advanced age. Meanwhile, as expected from processing speed theory, we observed that the overall picture naming speed and accuracy linearly decreased with age beyond the psychomotor slowing, replicating the widely established age-related decline in word-finding ability (A. S. Brown & Nix, 1996; D. M. Burke et al., 1991, 2004; Maylor, 1990, 1997; T. A. Salthouse & Mandell, 2013). With CI indexing the potential inhibition deficit that happens at the stage of lexical selection (Abdel Rahman & Melinger, 2011; Belke et al., 2005; Damian et al., 2001; Schnur et al., 2006), we found that older adults experience increased CI, reflecting an increased interference from competing distractors. This result was aligned with the prediction of both the inhibition and transmission deficit hypotheses.

On the other hand, changes in the degree of PF across age can inform the age-related alterations that occur during phonological activation (de Zubicaray & McMahon, 2009; de Zubicaray et al., 2002; Meyer, 1996; Meyer & Schriefers, 1991; Roelofs, 1992; Schriefers et al., 1990). Our analysis showed that the degree of PF in naming RT decreased with age. That is, older adults benefited less from phonological primes compared to young adults in the speed of word retrieval. This result is inconsistent with the inhibition deficit hypothesis's prediction that PF would increase across age due to older adults' increased susceptibility to distractors (Healey et al., 2008). However, a PF increase with age may be intrinsically difficult to observe since younger adults likely already perform at ceiling during primed picture naming, leaving less room for improvement across age. As observed in the picture naming RT and accuracy composite score (see Supplementary 2.2), the degree of PF remained stable across age. Although the PF trajectories across age yielded a rather inconclusive argument against the inhibition deficit hypothesis, they can be explained within the framework of transmission deficit hypothesis. The PF decrease with age is consistent with the assertion that older adults have a weaker connection between nodes; therefore, despite the bottom-up converging activation of the phonological primes being less vulnerable to age-related deterioration, older adults still have a less efficient top-down lemma to phonology activation process (MacKay & Burke, 1990). Moreover, the age-invariant PF in the composite score (see Supplementary 2.2) replicated previous studies supporting the transmission deficit hypothesis that top-down lemma to phonology activation is resistant to the effect of aging (i.e., Taylor & Burke, 2002). Summarizing the first step of the study, processing speed theory was supported by the overall picture naming speed and accuracy decrease across age. Transmission deficit hypothesis was supported by both the age-related CI increase and the PF decrease during lemma selection and phonological activation, respectively. Meanwhile, although the CI increase with age also conformed to the inhibition deficit hypothesis, the PF decrease with age contradicted the inhibition deficit prediction.

4.2. Connected speech disfluency explained picture naming speed but not accuracy, CI, or PF

If distractor activation during lemma selection, or the top-down connection from lemma to phonology explained observed word-finding difficulties, as the transmission deficit hypothesis would argue, then the degree of CI and PF should be associated with measurements of word-finding difficulties. In the recent decades, one increasingly popular way of measuring word-finding difficulties in healthy and pathological aging is connected speech characterization (Bortfeld et al., 2001; Castro & James, 2014; Horton et al., 2010; Slegers et al., 2018). Studies on aphasic patients have also related these natural speech disfluencies with picture naming performance (see Mason & Nickels, 2022 for a review) and clinician ratings of word-finding difficulties (Yeung et al., 2021). The current study explored the associations between the CI/PF effects in the picture naming game performance and speech disfluencies measured in natural speech across young and older adults. The results suggested that although the CI and PF trajectories with age supported the transmission deficit hypothesis, no significant association between CI/PF and natural speech disfluency was observed. On the other hand, a robust relationship between overall picture naming speed and disfluency in natural speech existed. Our analysis showed that faster picture naming RT was associated with less speech disfluency, especially in the aspects related to the speed of speech (e.g., speech rate and average word duration).

Although this association is seemingly unsurprising, previous studies investigating the relations between single-word confrontational naming and word-finding features in natural speech in healthy and aphasic adults concluded that these two types of word-finding measures (single-word retrieval vs. connected speech) do not measure the same word-finding mechanism (see Kavé & Goral, 2017 for a review). Clearly, there are major differences between single word retrieval and connected speech, in part because compensatory strategies such as substitution words and circumlocutions can be used during connected speech when experiencing word-finding difficulties, and there is not a particular correct answer in connected speech, as there is in a single-word retrieval task. However, most previous studies only measured single-word confrontational naming *accuracy* when examining its relation to connected speech features. Interestingly, the findings in the current study suggest the possibility that single-word retrieval RT, but not accuracy, allows us to tap into the shared mechanism behind age-related word-finding difficulty measured in single-word production and connected speech. Although picture naming accuracy was likely to be at ceiling for healthy adults, there was still a robust positive correlation between faster covert picture naming and faster, more fluent natural speech, reflecting a well-maintained word-finding ability. Hence, the association that exists between the game measures and word-finding in natural speech seems to involve overall speed of picture naming rather than naming accuracy, or the level of CI and PF.

4.3. Both verbal and non-verbal executive factors associated with PWI performance

Based on our results showing age-related changes in picture naming performance and their relationship with disfluency in natural speech, processing speed theory seems to

play an important role in age-related decline in word-finding ability. Nonetheless, since processing speed theory is an information-universal theory, associations between the domain-specific picture naming ability and general executive function need to be further investigated. In the third step of the study, we collected a set of executive measures and associated them with picture naming performance in the PWI game, to investigate the relationship between executive function and word-retrieval ability. As expected, all executive functions declined with age. When we extracted factors from the executive measures, all the verbal fluency, switching, and WM related measures loaded onto the first factor. For the second factor, inhibition (perceptual and motor inhibition) measures loaded together. Thus, the factor analysis separated the verbal executive tasks from the non-verbal measures. When predicting picture naming accuracy with executive factors, the interaction between game condition and executive factor 1 indicated that verbal fluency, switching, and WM ability were associated with the amount of CI during picture naming. In other words, the less fluent someone is in generating and switching between words with phonological or categorical relations, the more interference from competing words she experiences during word retrieval. This suggests that people who experience more interference between closely connected words in their mental lexicon will show slower RT and lower accuracy in tasks that require smooth navigation between words. The opposing relationship between inhibition and shifting ability has also been observed in a large number of studies, suggesting a trade-off between the stability (strong maintenance of goal) and flexibility (switching between goals) aspects of executive function (see Friedman & Miyake, 2017).

On the other hand, an interaction between game condition and the second executive factor (motor and perceptual inhibition), in predicting picture naming RT was also observed, suggesting that better inhibition ability corresponds to a stronger picture naming facilitation from phonological primes. This association found between word retrieval and general cognitive ability suggests that the inhibitory factor may explain the decline in phonological activation efficiency, showing some evidence of a shared mechanism behind word retrieval and non-verbal cognitive control. Furthermore, we observed that not only verbal fluency, switching, and WM, but also motor and perceptual inhibition, helped predict the overall picture naming speed and accuracy. The more impaired the executive functioning (in all measured cognitive aspects), the more slowly and less accurately participants retrieved picture names. These results support the information universal account that all domains of cognitive ability are likely linked to a common mechanism where when one aspect deteriorates, other aspects deteriorate as well.

4.4. Limitations

One potential limitation of the study concerns the design of the PWI task. The participants were required to covertly name pictures while monitoring the rule of responding based on the ending sound. This additional retention of the response rule may have made the PWI task somewhat of a "dual task," given that the difficulty of dual processing has its unique interaction with age during language production (Kemper et al., 2009, 2011). It is possible that the age-related decrease in PF reflects the age-related difference in dual-processing where young adults can benefit more from phonological primes due to their

higher ability to dual-process compared to older adults. We cannot completely account and control for the additional cognitive load imposed by this design of the PWI task in our analysis. Nonetheless, it is important to point out that the task design replicated the latency difference between the optimal elicitation of CI and PF (Wei et al., 2022), suggesting that the measured effects match the timeline theorized for the distinct stages of word retrieval during overt naming. Moreover, unlike most of the dual processing studies in language production where participants are required to monitor two distinct cognitive processes (e.g., a language production task paired with a visuo-motor task) (Kemper et al., 2009, 2011), the ending sound matching design in our PWI task arguably focuses on only one cognitive task at a given moment. Participants were not required to memorize the target sound of the block since it was always shown on the screen. When they completely activated the phonologies of the picture names, they naturally ended the picture naming mental process on the ending sound of the name, forming the basis of their manual response. Future studies with a simpler design of overt picture naming may be beneficial to ensure that the age-related effects observed in the PWI task exist without the ending sound judgment design. Nevertheless, the current design of the PWI has proven quite convenient for a large-scale online investigation of word production processes across the lifespan, together with measurements of natural speech and executive function.

5. Conclusion

Summarizing the three steps of investigation in this study, CI and PF trajectories across age supported the transmission deficit hypothesis but not the inhibition deficit hypothesis, while the overall picture naming speech and accuracy aligned with the processing speed theory. However, no association between CI/PF and word-finding disfluencies in natural speech was observed, limiting the support for the transmission deficit hypothesis. On the other hand, a robust association was found between picture naming RT in the game and natural speech disfluency, highlighting the role of processing speed in revealing word-finding struggles in both single-word and in-context word retrieval. Moreover, picture naming RT and accuracy were associated with both the verbal fluency, switching, and WM factor and the (non-verbal) motor and perceptual inhibition factor, while CI was related to the verbal executive factor and PF to the non-verbal inhibition factor. Overall, the findings of this study suggest that (1) more evidence was revealed for the information-universal than the information-specific angle of aging theory, and that (2) the common mechanism behind age-related decline in word finding and executive control, if it truly exists, is more likely to be cognitive slowing (T. Salthouse, 1996) rather than deficits in cognitive inhibition (Zacks & Hasher, 1994). Theoretically speaking, this study suggested the possibility that the processing speed theory could supersede inhibition and transmission deficit hypotheses in not only explaining but also measuring age-related word-finding difficulty. In the application of this research, the present study indicates that the most sensitive age-related word-finding difficulty measure is likely picture naming RT instead of accuracy, even though the latter is usually the focus of traditional naming measures implemented on standard cognitive exams such as the MoCA (Nasreddine et al., 2005) (see Supplementary 3.1 for MoCA's bias toward naming accuracy in our dataset). As suggested by the theoretical implications and the empirical findings of

this study, if processing speed lies at the center of cognitive decline that can behaviorally reflect as word-finding difficulty, it is crucially important to incorporate picture naming RT into standard cognitive assessments to facilitate earlier detection of progression into pathological aging.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by an internship grant FR75766 from the Mitacs Accelerate program to HTW, NSERC Discovery Grant RGPIN-2019-06515 to JAM, and a Connaught Innovation Award IA-2021-22 to MHC.

ORCID

Hsi T. Wei http://orcid.org/0000-0003-4852-8288
Jed A. Meltzer http://orcid.org/0000-0002-4301-1901

Data availability statement

All data, materials, codes, and supplementary materials, behind this analysis have been made publicly available on Open Science Framework (https://osf.io/yc7jm/?view_only=40bbaf6d13314b0dbee81ce41b0aa757).

References

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 149–161. https://doi.org/10.1037/a0021208

Abel, S., Dressel, K., Bitzer, R., Kummerer, D., Mader, I., Weiller, C., & Huber, W. (2009). The separation of processing stages in a lexical interference fMRI-paradigm. *Neuroimage: Reports*, *44*(3), 1113–1124. https://doi.org/10.1016/j.neuroimage.2008.10.018

Alario, F. X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *53*(3), 741–764. https://doi.org/10.1080/027249800410535

Amrhein, P. C. (1995). Evidence for task specificity in age-related slowing: A review of speeded picture-word processing studies. In P. A. Allen & T. R. Bashore (Eds.), *Advances in psychology* (Vol. 110, pp. 143–170). North-Holland. https://doi.org/10.1016/S0166-4115(06)80069-5

Arbuckle, T. Y., Nohara LeClair, M., & Pushkar, D. (2000). Effect of off-target verbosity on communication efficiency in a referential communication task. *Psychology and Aging*, *15*(1), 65–77. https://doi.org/10.1037/0882-7974.15.1.65

Beier, E. J., Chantavarin, S., & Ferreira, F. (2023). Do disfluencies increase with age? Evidence from a sequential corpus study of disfluencies. *Psychology and Aging*, *38*(3), 203–218. https://doi.org/10.1037/pag0000741

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology A, Human Experimental Psychology*, *58*(4), 667–692. https://doi.org/10.1080/02724980443000142

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, *44*(2), 123–147. https://doi.org/10.1177/00238309010440020101

Brehman, D., Burns, S., Thaler, N., Rojas, S., & Barchard, K. (2009). *Fatigue: A threat to study results*.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

Brown, A. S., & Nix, L. A. (1996). Age-related changes in the tip-of-the-tongue experience. *The American Journal of Psychology*, *109*(1), 79–91. https://doi.org/10.2307/1422928

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 325–337. https://doi.org/10.1016/S0022-5371(66)80040-3

Burke, D., Mackay, D., & James, L. (2012). Beneficial effects of novel antagonists of GHRH in different models of Alzheimer's disease. *Models of Cognitive Aging*, *4*(11), 755–767. https://doi.org/10.18632/aging.100504

Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. *Psychological Science*, *15*(3), 164–170. https://doi.org/10.1111/j.0956-7976.2004.01503004.x

Burke, D. M., Mackay, D. G., Evans, J. G., Holliday, R., Kirkwood, T. B. L., Laslett, P., & Tyler, L. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *352*(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, *30*(5), 542–579. https://doi.org/10.1016/0749-596X(91)90026-G

Castro, N., & James, L. E. (2014). Differences between young and older adults' spoken language production in descriptions of negative versus neutral pictures. *Aging, Neuropsychology & Cognition*, *21*(2), 222–238. https://doi.org/10.1080/13825585.2013.804902

Cohen, G. (1994). Age-related problems in the use of proper names in communication. In *Interpersonal communication in older adulthood: Interdisciplinary theory and research* (pp. 40–57). Sage Publications, Inc. https://doi.org/10.4135/9781483326832.n3

Curcic, J., Vallejo, V., Sorinas, J., Sverdlov, O., Praestgaard, J., Piksa, M., Deurinck, M., Erdemli, G., Bügler, M., Tarnanas, I., Taptiklis, N., Cormack, F., Anker, R., Massé, F., Souillard-Mandar, W., Intrator, N., Molcho, L., Madero, E., Bott, N., & Hannesdottir, K. (2022). Description of the method for evaluating digital endpoints in Alzheimer disease study: Protocol for an exploratory, cross-sectional study. *JMIR Research Protocols*, *11*(8), e35442. https://doi.org/10.2196/35442

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), B77–B86. https://doi.org/10.1016/S0010-0277(01)00135-4

de Zubicaray, G. I., & McMahon, K. L. (2009). Auditory context effects in picture naming investigated with event-related fMRI. *Cognitive, Affective & Behavioral Neuroscience*, *9*(3), 260–269. https://doi.org/10.3758/cabn.9.3.260

de Zubicaray, G. I., McMahon, K. L., Eastburn, M. M., & Wilson, S. J. (2002). Orthographic/Phonological facilitation of naming responses in the picture-word task: An event-related fMRI study using overt vocal responding. *Neuroimage: Reports*, *16*(4), 1084–1093. https://doi.org/10.1006/nimg.2002.1135

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. Psychological Corp.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dennis, P. A., & Hess, T. M. (2016). Aging-related gains and losses associated with word production in connected speech. *Neuropsychology, Development, and Cognition Section B, Aging, Neuropsychology and Cognition*, *23*(6), 638–650. https://doi.org/10.1080/13825585.2016.1158233

Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, *13*(1), 121–143. https://doi.org/10.1162/089892901564117

Faber-Langendoen, K., Morris, J. C., Knesevich, J. W., LaBarge, E., Miller, J. P., & Berg, L. (1988). Aphasia in senile dementia of the Alzheimer type. *Annals of Neurology*, *23*(4), 365–370. https://doi.org/10.1002/ana.410230409

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Filiou, R.-P., Bier, N., Slegers, A., Houzé, B., Belchior, P., & Brambati, S. M. (2020). Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment: A scoping review. *Aphasiology*, *34*(6), 723–755. https://doi.org/10.1080/02687038.2019.1608502

Fine, E. M., & Delis, D. C. (2011). Delis–Kaplan executive functioning system. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 796–801). Springer. https://doi.org/10.1007/978-0-387-79948-3_1539

Folstein, M., Folstein, S., & McHugh, P. (1975). "Mini-mental state" a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Fraser, K. C., Meltzer, J. A., Rudzicz, F., & Garrard, P. (2015). Linguistic features identify Alzheimer's disease in narrative speech. *Journal of Alzheimer's Disease*, *49*(2), 407–422. https://doi.org/10.3233/jad-150520

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, *37*(4), 581–590. https://doi.org/10.3758/BF03192728

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

Germain, S., & Collette, F. (2008). Dissociation of perceptual and motor inhibitory processes in young and elderly participants using the Simon task. *Journal of the International Neuropsychological Society*, *14*(6), 1014–1021. https://doi.org/10.1017/S135561770808123X

Goodglass, H., & Kaplan, E. (1983). *The Boston diagnostic aphasia examination*.

Hale, S., & Myerson, J. (1996). Experimental evidence for differential slowing in the lexical and nonlexical domains. *Aging, Neuropsychology & Cognition*, *3*(2), 154–165. https://doi.org/10.1080/13825589608256621

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The Psychology of Learning and Motivation: Advances in Research and Theory*, *22,* 193–225.

Healey, M. K., Campbell, K. L., & Hasher, L. (2008). Cognitive aging and increased distractibility: Costs and potential benefits. *Progress in Brain Research*, *169*, 353–363. https://doi.org/10.1016/S0079-6123(07)00022-2

Higby, E., Cahana-Amitay, D., Vogel-Eyny, A., III, Albert, M. L., & Obler, L. K. (2019). The role of executive functions in object- and action-naming among older adults. *Experimental Aging Research*, *45*(4), 306–330. https://doi.org/10.1080/0361073X.2019.1627492

Hillis, A. E. (2013). *The handbook of adult language disorders*. Psychology Press.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30* (2), 179–185. https://doi.org/10.1007/BF02289447

Horton, W. S., Spieler, D. H., & Shriberg, E. (2010). A corpus analysis of patterns of age-related change in conversational speech. *Psychology and Aging*, *25*(3), 708–713. https://doi.org/10.1037/a0019424

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*(255). https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology Learning, Memory, and Cognition*, *26*(6), 1378–1391. https://doi.org/10.1037/0278-7393.26.6.1378

Jennings, J. R., Mendelson, D. N., Redfern, M. S., & Nebes, R. D. (2011). Detecting age differences in inhibition processes with a test of perceptual and motor inhibition. *Experimental Aging Research*, *37*(2), 179–197. https://doi.org/10.1080/0361073X.2011.554512

Jescheniak, J. D., & Schriefers, H. (2001). Priming effects from phonologically related distractors in picture—word interference. *The Quarterly Journal of Experimental Psychology Section A*, *54*(2), 371–382. https://doi.org/10.1080/713755981

Kavé, G., & Goral, M. (2017). Do age-related word retrieval difficulties appear (or disappear) in connected speech? *neuropsychology, development, and cognition*. *Section B: Aging, Neuropsychology and Cognition*, *24*(5), 508–527. https://doi.org/10.1080/13825585.2016.1226249

Kemper, S., Schmalzried, R., Herman, R., Leedahl, S., & Mohankumar, D. (2009). The effects of aging and dual task demands on language production. *Neuropsychology, Development, and Cognition Section B, Aging, Neuropsychology and Cognition*, *16*(3), 241–259. https://doi.org/10.1080/13825580802438868

Kemper, S., Schmalzried, R., Herman, R., & Mohankumar, D. (2011). The effects of varying task priorities on language production by young and older adults. *Experimental Aging Research*, *37*(2), 198–219. https://doi.org/10.1080/0361073X.2011.554513

Kim, S., Hasher, L., & Zacks, R. T. (2007). Aging and a benefit of distractability. *Psychonomic Bulletin & Review*, *14*(2), 301–305. https://doi.org/10.3758/BF03194068

Kirshner, H. S. (2012). Primary progressive aphasia and Alzheimer's disease: Brief history, recent evidence. *Current Neurology and Neuroscience Reports*, *12*(6), 709–714. https://doi.org/10.1007/s11910-012-0307-2

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, *8*(1), 34–43. https://doi.org/10.1037/0882-7974.8.1.34

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Lima, S. D., Hale, S., & Myerson, J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging*, *6*(3), 416–425. https://doi.org/10.1037/0882-7974.6.3.416

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, *130*(2), 199–207. https://doi.org/10.1037/0096-3445.130.2.199

Mackay, D. G. (1988). The organization of perception and action. A theory for language and other cognitive skills. *The Italian Journal of Neurological Sciences*, *9*(3), 303–303. https://doi.org/10.1007/BF02334060

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In F. G. A. Stone & R. West (Eds.), *Advances in psychology* (Vol. 71, pp. 213–263). North-Holland. https://doi.org/10.1016/S0166-4115(08)60159-4

Madden, D., & Allen, P. A. (2015). History of cognitive slowing theory and research. In N. A. Pachana (Ed.), *Encyclopedia of Geropsychology* (pp. 1–10). Springer. https://doi.org/10.1007/978-981-287-080-3_96-1

Mason, C., & Nickels, L. (2022). Are single-word picture naming assessments a valid measure of word retrieval in connected speech? *International Journal of Speech-Language Pathology*, *24*(1), 97–109. https://doi.org/10.1080/17549507.2021.1966098

May, C. P. (1999). Synchrony effects in cognition: The costs and a benefit. *Psychonomic Bulletin & Review*, *6*(1), 142–147. https://doi.org/10.3758/bf03210822

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition*, *27*(5), 759–767. https://doi.org/10.3758/bf03198529

Mayer, J., & Murray, L. (2003). Functional measures of naming in aphasia: Word retrieval in confrontation naming versus connected speech. *Aphasiology*, *17*(5), 481–497. https://doi.org/10.1080/02687030344000148

Maylor, E. A. (1990). Recognizing and naming faces: Aging, memory retrieval, and the tip of the tongue state. *Journal of Gerontology*, *45*(6), 215–P226. https://doi.org/10.1093/geronj/45.6.P215

Maylor, E. A. (1997). Proper name retrieval in old age: Converging evidence against disproportionate impairment. *Aging, Neuropsychology & Cognition*, *4*(3), 211–226. https://doi.org/10.1080/13825589708256648

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language*, *35*(4), 477–496. https://doi.org/10.1006/jmla.1996.0026

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(6), 1146. https://doi.org/10.1037//0278-7393.17.6.1146

Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *Journal of Clinical and Experimental Neuropsychology*, *40*(9), 917–939. https://doi.org/10.1080/13803395.2018.1446513

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nassauer, K., & Halperin, J. (2003). Dissociation of perceptual and motor inhibition processes through the use of novel computerized conflict tasks. *Journal of the International Neuropsychological Society*, *9*(1), 25–30. https://doi.org/10.1017/S1355617703910034

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, *97*(1), 593–608. https://doi.org/10.1121/1.412282

Robin, J., Xu, M., Kaufman, L. D., & Simpson, W. (2021). Using digital speech assessments to detect early signs of cognitive impairment. *Frontiers in Digital Health*, *3*, 3. https://doi.org/10.3389/fdgth.2021.749758

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the stroop task. *Psychological Review*, *110*(1), 88–125. https://doi.org/10.1037/0033-295X.110.1.88

Roelofs, A. (2008). Attention to spoken word planning: Chronometric and neuroimaging evidence. *Language and Linguistics Compass*, *2*(3), 389–405. https://doi.org/10.1111/j.1749-818X.2008.00060.x

Rohrer, J. D., Knight, W. D., Warren, J. E., Fox, N. C., Rossor, M. N., & Warren, J. D. (2008). Word-finding difficulty: A clinical analysis of the progressive aphasias. *Brain A Journal of Neurology*, *131*(1), 8–38. https://doi.org/10.1093/brain/awm251

Rowe, G., Valderrama, S., Hasher, L., & Lenartowicz, A. (2006). Attentional Disregulation: A Benefit for Implicit Memory. *Psychology and Aging*, *21*(4), 826–830. https://doi.org/10.1037/0882-7974.21.4.826

Salthouse, T. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Salthouse, T. A. (2001). Structural models of the relations between age and measures of cognitive functioning. *Intelligence*, *29*(2), 93–115. https://doi.org/10.1016/S0160-2896(00)00040-4

Salthouse, T. A. (2017). Shared and unique influences on age-related cognitive change. *Neuropsychology*, *31*(1), 11–19. https://doi.org/10.1037/neu0000330

Salthouse, T. A., & Mandell, A. R. (2013). Do age-related increases in tip-of-the-tongue experiences signify episodic memory impairments? *Psychological Science*, *24*(12), 2489–2497. https://doi.org/10.1177/0956797613495881

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences*, *50*(6), 297–306. https://doi.org/10.1093/geronb/50b.6.p297

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227. https://doi.org/10.1016/j.jml.2005.10.002

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*(1), 86–102. https://doi.org/10.1016/0749-596X(90)90011-N

Semmelmann, K., & Weigelt, S. (2017). Online psychophysics: Reaction time effects in cognitive experiments. *Behavior Research Methods*, *49*(4), 1241–1260. https://doi.org/10.3758/s13428-016-0783-4

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*, 5. https://doi.org/10.3389/fpsyg.2014.00772

Slegers, A., Filiou, R.-P., Montembeault, M., & Brambati, S. M. (2018). Connected speech features from picture description in Alzheimer's disease: A systematic review. *Journal of Alzheimer's Disease: JAD*, *65*(2), 519–542. https://doi.org/10.3233/JAD-170881

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, *11*(2), 333–341. https://doi.org/10.1037/0882-7974.11.2.333

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, *14*(3), 458–472. https://doi.org/10.1037/0882-7974.14.3.458

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psychology and Aging*, *17*(4), 662–676. https://doi.org/10.1037/0882-7974.17.4.662

Thornton, R., & Light, L. L. (2006). Language comprehension and production in normal aging. In J. E. Birren, K. W. Schaie, R. P. Abeles, M. Gatz, & T. A. Salthouse (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 261–287). Academic Press. https://doi.org/10.1016/B978-012101264-9/50015-X

Townsend, J., & Ashby, F. (1983). *The stochastic modeling of elementary psychological processes*. *Part 2*.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In *Cognitive theory*. Psychology Press.

Tun, P. A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging*, *17*(3), 453–467. https://doi.org/10.1037/0882-7974.17.3.453

Wei, H. T., Hu, Y. Z., Chignell, M., & Meltzer, J. A. (2022). Picture-word interference effects are robust with covert retrieval, with and without gamification. *Frontiers in Psychology*, *12*, 12. https://doi.org/10.3389/fpsyg.2021.825020

Wilshire, C., Singh, S., & Tattersall, C. (2016). Serial order in word form retrieval: New insights from the auditory picture–word interference task. *Psychonomic Bulletin & Review*, *23*(1), 299–305. https://doi.org/10.3758/s13423-015-0882-8

Yeung, A., Iaboni, A., Rochon, E., Lavoie, M., Santiago, C., Yancheva, M., Novikova, J., Xu, M., Robin, J., Kaufman, L. D., & Mostafa, F. (2021). Correlating natural language processing and automated speech analysis with clinician assessment to quantify speech-language changes in mild cognitive impairment and Alzheimer's dementia. *Alzheimer's Research & Therapy*, *13*, 109. https://doi.org/10.1186/s13195-021-00848-x

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In *Inhibitory processes in attention, memory, and language* (pp. 241–264). Academic Press.

Zemla, J. C. (2022). Knowledge representations derived from semantic fluency data. *Frontiers in Psychology*, *13*, 815860. https://doi.org/10.3389/fpsyg.2022.815860

View publication stats