The Trait-Like Nature of Bridging and Instrument Inferences in Younger and Older Adults: An Individual Differences Study

Sonny R. Wang & Craig G. Chambers
University of Toronto

There has been growing recognition in psycholinguistics that exploring individual differences (IDs) in language processing can inform key theoretical debates [1] and elucidate the influence of experience or domain-general cognition [2,3]. However, many studies assessing correlations between performance on linguistic tasks and general cognitive abilities show varying and contradictory results [4,5], which may stem from low internal consistency in the linguistic stimuli [6,7]. Further, few studies exclude the possibility that IDs could result from spurious differences in participants' mental state at the **moment of testing** (e.g., mood, fatigue, confidence), which again could make the measures unreliable. Here, we report a large-scale study intended to bypass these limitations by: (i) comparing (English-speaking) participants' performance on different linguistic inferencing tasks **using precisely the same measure** (to explore whether the tasks draw on similar mental resources), (ii) testing participants at **two time points** 4 weeks apart (to determine test-retest reliability, ensuring IDs are indeed stable/trait-like), and (iii) exploring how these relate to **age differences** (younger vs. older adults) to assess whether reported greater **intra-individual** variability in older adults' cognitive performance [8,9] would make stable linguistic performance across time and tasks less likely for that age group.

In **Expts 1** (*N*=52) and **2** (*N*=54), we validated our self-paced sentence-by-sentence reading materials measuring instrument and bridging inferences, respectively. Critical passages contained target sentences that were either congruent or incongruent with instrument- or bridging inferences that could be drawn upon reading earlier sentences (see **Table 1**). Results replicated previous findings [10,11], showing readers readily make instrument/bridging inferences during reading (longer avg. residualized RTs for incongruent vs. congruent target sentences). A subset of the materials was then used in a **large-scale, pre-registered longitudinal and correlational study** for both **younger adults** (**Expt 3**, age 18-38, *N*=150), and **older adults** (**Expt 4**, age 55-75, *N*=150). The first and second sessions (separated by 4 weeks) involved the same 24 critical passages (12 for bridging; 12 for instrument inferences) but different fillers. Results replicated the condition differences in Expts 1&2 (see **Fig. 1**). Next, we calculated an individual score per participant, (difference in averaged residualized reading time for incongruent minus congruent target sentences, collapsed across passages). We then subjected the rankings of individual scores from different age, task, and time points to a series of **correlational analyses** to assess whether the two inference types draw on similar resources, and whether IDs are stable over time (see **Table 2**, note the same patterns were found using raw [unranked] composite scores, which are less stringent). Results showed significant within-participant correlations in performance **between the two inference tasks,** consistent with the idea that bridging and instrument inferences draw on overlapping resources. Further, IDs in performance on both tasks **were stable across 4 weeks**, strongly supporting the idea that the resources underlying inferencing are trait-like and not the product of a given moment of testing. Finally, these patterns were found for **both younger and older adults**, further reinforcing the intact nature of inferencing abilities across the adult lifespan. We also report a high average between-participant internal consistency for our stimuli used in both tasks for both ages and time points (Cronbach's α = 0.88). In sum, the work carries both methodological and theoretical implications: The materials used in the present study demonstrated high internal consistency and test-retest reliability in capturing individual differences in linguistic inferencing. Further, the results provided clear evidence for shared mechanisms across different inference types, and showed the extraordinary stability of inferencing abilities across the adult lifespan.

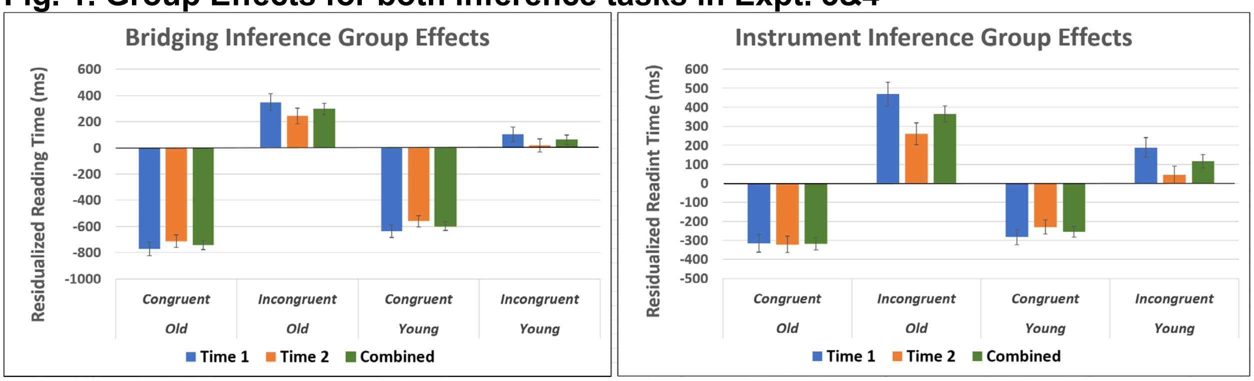
References. [1] Kidd et al., 2018. *Trends in Cog Sciences*. [2] Farmer et al., 2017. *Quarterly Jnl of Expt'l Psych*. [3] Olknoniemi e al., 2016. *Jnl of Expt'l Psych.* [4] Ryskin et al., 2015. *Jnl of Expt'l Psych*. [5] Arnon, 2020. *Behavioral Research Methods*. [6] Cunnings & Fujita, 2021. *Applied Psycholinguistics.* [7] James et al., 2018. *Journal of Memory & Language*. [8] Shammi eta al., 1998. *Aging, Neuropsychology and Cognition.* [9] Fagot et al., 2018. *Jnl of Intelligence.* [10] Singer et al., 1992. *Memory & Cognition.* [11] Vukic et al., 2009. *Disc. Processes*.

Table 1. Example Materials

Instrument Inference (**Expts 1,3,4.** Note: there are 5 sentences in each passage, not all are shown)	
Antecedent	Dan was building a new shed he had been planning all winter. For the past ten minutes, he had been pounding nails into the boards. *Instrument Inference -> [Dan is using a] hammer*
Target (congruent)	"Lucky I got a hammer!" Dan thought to himself.
Target (incongruent)	"If only I had a hammer!" Dan thought to himself.

Bridging Inference (**Expt 2,3,4.** Note: there are 5 sentences in each passage, not all are shown)	
Antecedent	Carol went outside and had forgotten about the bread she was baking in the oven. (Inference –> oven is still on; bread is still inside)
Target (congruent)	When she finally came back inside, Carol saw smoke coming from the oven.
Target (incongruent)	When she finally came back inside, Carol decided to turn on the oven.

Fig. 1. Group Effects for both inference tasks in Expt. 3&4



All studies were conducted online using Gorilla. Data screening and preparation steps: [i] make sure all participants have full number of observations; [ii] eliminate inattentive participants; [iii] calculate residualized reading times by participant to control for sentence length effects and intra-individual differences in reading speed.

LME

Model:	residualized reading time ~ congruency * age * task
Intercepts:	participant + passage
Slopes:	(congruency x participant) + (task x participant) + (congruency x task x participant) + (age x passage)

Analyses for the inferencing tasks used LME models with a maximal random effects structure.

Table 2. Correlations based on rankings of participants' individual scores

	Correlations between Bridging and Instrument Inference Tasks						Test-Retest Correlation (Time 1 ~ Time 2)			
	Time 1		Time 2		Combined Times		BRI		INS	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Young	0.22	**	0.16	0.06	0.20	***	0.24	**	0.27	***
Old	0.30	***	0.45	***	0.40	***	0.34	***	0.35	***
Combined Age	0.31	***	0.37	***	0.35	***	0.36	***	0.35	***