Testing the portability of outcomes across eye-tracking and self-paced reading reveals that traditional assumptions about spillover in relative clause sentences do not hold

Eye-tracking and self-paced reading (SPR) are popular methods of studying online sentence processing. Psycholinguists often compare findings across these methods or predict how findings in one method would manifest based on findings in the other. For example, when comparing relative clause (RC) effects in eye-tracking and SPR, psycholinguists often predict spillover, where they expect RC effects observed in eye-tracking to emerge in later regions in SPR because the manual demands involved in SPR delay the emergence of the RC effect[1] (Fig 1). In other words, psycholinguists expect that the results they observe when participants read RC sentences in eye-tracking would directly map onto another set of results (i.e., spillover) *had the same participants read the same sentences in SPR instead* and vice versa. This expectation assumes that eye-tracking and SPR as methods of reading are ***portable***: i.e., that we can directly link a sample's (hypothetical) outcomes from both methods, enabling us to make comparisons and predictions across them. Researchers often assume that the portability assumption holds, where they collect data using both methods as separate experiments, treat the samples from the experiments as exchangeable, and compare the results[1,2,3]. However, whether the portability assumption holds is an empirical question that has never been directly tested. We perform that direct test in this study by having participants read the same sentences in both methods.

224+ participants read the same 120 English sentences in the same order in separate (counterbalanced) eye-tracking and SPR sessions that were 7+ days apart (M = 11.87*; SD* = 7.06). Eighty sentences included a subject/object-extracted RC (SRC/ORC) manipulation. For analysis, we divided the sentences into regions as done in previous work[4,5] (Fig 1).

We first examined overall portability by correlating average eye-tracking outcomes (i.e., gaze duration, go-past duration, total duration) with average SPR RTs. Naturally, we expect strong correlations *within* a region (i.e., the same words are read), but if spillover holds, we expect eye-tracking outcomes in earlier regions to also be strongly correlated with SPR outcomes in *later* regions (Fig 2). After adjusting for overall processing speed, people's average eye-tracking outcomes do not correlate with their SPR outcomes (Fig 3a). In contrast, average eye-tracking outcomes for an item correlate with the average SPR outcomes for that item but little in the way predicted by spillover. A standout pattern is that when more time is spent in earlier regions in eye-tracking, the less time is spent in those same regions in SPR and vice versa (Fig 3b, blue bands of negative correlations on right and bottom edge of matrices; r = [-0.28, -0.59]), suggesting that items are eliciting different reading strategies when they're read in different methods.

We then examined the portability of the RC effect itself across both methods and tested spillover by fitting multivariate Bayesian mixed-effect models, which allow the simultaneous modeling of both eye-tracking and SPR outcomes and account for the covariation between the outcomes due to being obtained from the same people and items. The overall RC effect was present in both eye-tracking and SPR although in different regions: it was only present in the RC in eye-tracking but present in both the RC and PP in SPR (Fig 4). This finding is traditionally taken as the evidence for spillover. However, the typical explanation of how spillover occurs invokes within-person or within-item linking across methods (i.e., lines 7-10 above), which is not achieved by only comparing the overall effects found in the two methods. To achieve within-person or -item linking, we examined the correlations between people and item random slopes and found limited portability of the RC effect at the item-level in the RC & PP and limited evidence for spillover (Fig 5).

These results suggest that although comparing overall effects across methods may be warranted, portability across eye-tracking and SPR is not a given and that hypotheses about how outcomes in one method map onto the other, such as spillover, require a direct test of the portability assumption.

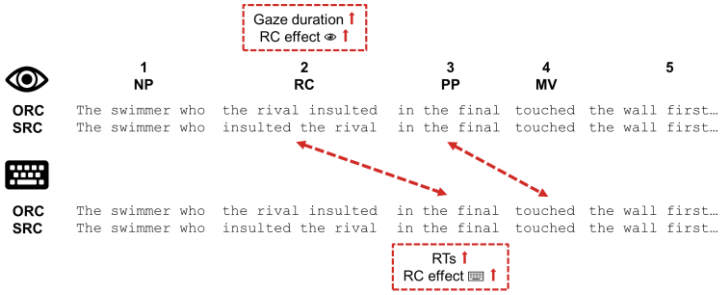

Figure 1. An illustration of the traditional spillover hypothesis for the RC effect: "RC effect spills over from RC in 👁 to PP/MV in ⌨ because of the manual demands involved in ⌨"

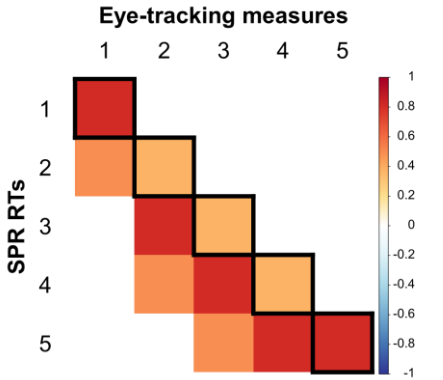

Figure 2. Correlation matrix of eye-tracking measures with SPR RTs. Any cross-section indicates the correlation between a region in eye-tracking (columns) and a region in SPR (rows). Figure shows the predictions for traditional spillover, where earlier regions in eye-tracking are strongly correlated with later regions in SPR.

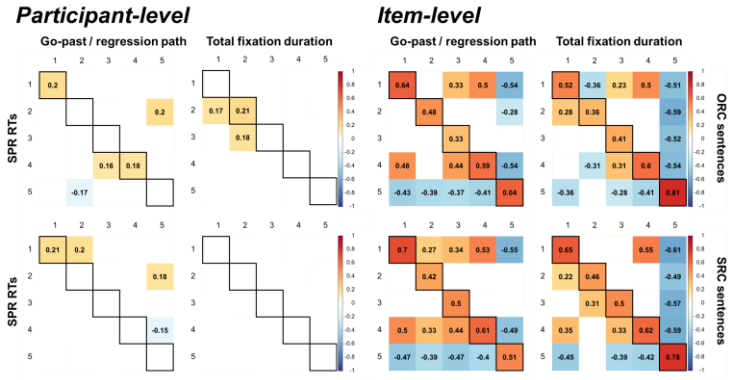

Figure 3. Results of overall portability analysis. Only go-past and total duration are shown here. (3A, left). After accounting for participants' overall processing speed, their eye-tracking measures are no longer correlated with their SPR RTs. (3B, right). After accounting for overall processing speed for items, the eye-tracking measures for items are still strongly correlated with SPR RTs.

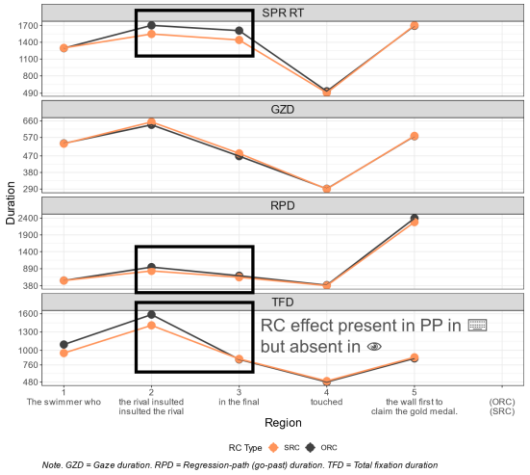

Figure 4. Overall RC effects observed in SPR RTs and various eye-tracking measures. Significant RC effect in the RC in go-past and total duration for eye-tracking, but significant RC effect in both the RC and PP in SPR, which is usually taken as evidence for spillover.

References:
[1]Roland, Mauner, & Hirose. (2021). *JML*, *119*. 104244. [2]Witzel, Witzel, & Forster (2012). *J Psychlx Res*, *41,* 105-128. [3]Kung. (2021). UNC Thesis. [4]Staub, Dillon, & Clifton. (2017). *CogSci*, *41*, 1353-1376. [5]Lowder & Gordon. (2021). *CogSci*, *45*, e13039.

Figure 5. Posterior means and 95% CI of correlations between RC effect parameters in eye-tracking and SPR within-person and within-item. Panels show model results for regions 2 (RC), 3 (PP), and 4 (MV), where each cross-section indicates the correlation between the RC effect at those regions for eye-tracking (columns) and SPR (rows). (5A, left) Correlations between participant RC effects in eye-tracking and SPR. Even within a region, there is no evidence that their RC effect in eye-tracking is correlated with their RC effect in SPR. (5B, right) Correlations between item RC effects in eye-tracking and SPR. RC effects in the PP across both methods are correlated. Hint of spillover from RC to PP for total duration.