



5.10.2021

Commentary report on Patrycja Kalamala Doctoral thesis Jagiellonian University “*Language experiences and cognitive control- Assessment and Interactions*” Krakow 2021

The thesis comprises an integrative synopsis, critique and envisaged future programme of research on the important and complex problem of language control in bilingual speakers together with three innovative and substantial studies (peer-reviewed and published) on which it is based with all materials, data and analysis scripts in open source. In this report, there is an in-depth understanding of the literature and the theoretical, methodological and statistical requirements for effective experimentation and advance. Below I consider matters raised in the synopsis and the three studies and mention matters that might be fruitfully considered in a *viva voce*. I have no hesitation in commending the award of a doctorate to Patrycja Kalamala subject to an anticipated satisfactory *viva voce*.

As the thesis makes clear language control in bilingual speakers is complex and I think the decision to concentrate on the role of inhibition as one putative process of cognitive control is well-motivated. Such a process is implicated in circumstances of conflict between alternative responses and is pertinent in bilingual speakers because of evidence consistent with, for example, the activation of lexical representations in both languages despite the ongoing use of just one. Studies 1 and 2 are guided by a theoretical proposal (the Adaptive Control Hypothesis, ACH) that argues that control processes adapt to the precise demands of language use. Some speakers, for example, may use their languages in distinct conversational contexts, others may switch between languages, depending on their interlocutor, within the same conversation (dual language use) though not within the same utterance. Conceivably, dual

language use increases demands on the process of response inhibition with regular demand leading to an increase in its efficiency. It was this possibility that led to a body of research examining the so-called bilingual advantage in which bilingual speakers seemed to show enhanced performance over speakers of just one language in suppressing a prepotent or dominant response (response inhibition). But if control processes adapt, then it is imperative to characterize how individuals within a sample of bilingual speakers actually use their languages and so establish possible boundary conditions on such an advantage and resolve issues of replicability and robustness.

Overall, the conception, design, sample power (N = 195 Polish-English participants) and analysis of Study 1 is exemplary. The study essentially sets the stage for the remaining two studies. Importantly, Study 1 sought to capture the psychometric validity of the questionnaire measures of language use unlike previous work by another group (Hartanto & Yang, 2016) in which this psychometric requirement had not been met as revealed by the reanalysis reported in Appendix B. Study 1's questionnaire data for the intensity of dual language use did meet this requirement. The co-occurrence of dual language use was captured as the distribution of language use with a more balanced use of two languages across a variety of contexts of language use indexed by a higher entropy score. The frequency of dual language use was captured in a measure of language mixing within utterances across the same set of contexts. A suitable summed index captured dual language context (DLC) intensity. [The study also reports use of a second measure of dual language use based on a questionnaire adapted from Hartanto & Yang, 2016 with convergent outcomes]. Experimental research offers a range of task that may tap response inhibition and so Study 1 made the principled decision to use multiple tasks (4) that purportedly tap this process and analysed these data using the latent variable approach. The study then assessed the association between DLC intensity and response inhibition using structural equation modelling. Unlike the questionnaire data, there was limited common variance in the performance of the response inhibition tasks despite their individual reliabilities. Indeed, structural equation modelling showed no association between DLC intensity and the response inhibition factor though indices of socio-economic status and Ravens IQ did yield evidence of an association despite the limited common variance of these response inhibition tasks. The issue of a possible

relation was pursued further by assessing association with the response inhibition measures considered separately with all null associations justified by Bayes factor analyses.

Despite the sample size, the population of bilinguals studied is one where English use is embedded in a Polish-speaking context. Subject to that sampling proviso, on the face of it, this experiment refutes a prediction derived from the ACH with respect to a control process involving response inhibition. However, as the thesis synopsis makes clear the study also raises questions as to whether or not any behavioural index of response inhibition or a behavioral index of any other control process postulated by the ACH will in fact be suitable to test the prediction because of the task impurity issue. Behavioral data convolve a number of distinct processes only some of which may be pertinent to test a specific prediction. Study 2 addresses this matter by adopting a methodology in which evoked reaction potential (ERP) data are collected alongside behavioral data. Despite the methodological advances of Study 1, it presumed that language control experiences would readily modulate performance on laboratory tasks. However, a study of this type can only yield an association between prior experience and task performance. In order to establish a causal relationship it is necessary to manipulate that experience more directly. Study 2 innovated a way to do this whilst effectively retaining the approach of Study 1 of examining the after-effects of the manipulation. Indeed, Study 2 addresses a potentially problematic feature of Study 1, namely that the language of instructions in that study, as I surmise, was Polish. Given that L2 English was not context relevant, performance on the response inhibition tasks required the recruitment of long-term after-effects of L2 language use. However, the duration and context-sensitivity of such effects is unknown.

Study 2 manipulated the immediate context of language using a set of time-limited language games in a within-subjects design that recruited a sample ($n = 32$) of Polish-English bilinguals in Poland with moderate to high English proficiency. The participants reported more frequent use of Polish (L1) in their everyday lives with English use in distinct contexts combined with language switching within utterances in that context. These data were derived from use of the Hartanto & Yang 2016 questionnaire though I was uncertain as to whether it was adapted and assessed as per Study 1. In each of the games (that varied in their

difficulty) an experimental participant had to describe a layout of depicted objects on their screen so that an experimental confederate could reconstruct their positioning in their own screen. The two-fold purpose was to emulate aspects of conversational language use during task-oriented dialogue and so achieve a degree of ecological validity and to manipulate the nature of the conversational exchange: single language context - exclusive use of L1 (Polish, the dominant language) and exclusive use of L2 (English, the non-dominant language and one less frequently used everyday) and a dual language context -switching language in alternate turns with two confederates, one of whom used English whilst the other used Polish. Following each game, participants were fitted with an EEG cap and completed two standard inhibition tasks: a manual stop-signal task and vocal Stroop task. Relative to the baseline L1 game, analyses of the behavioral data of these tasks revealed no effect of the induced context of language use. By contrast, standard ERP indices, P3 for the stop-signal task and N450 for the Stroop task indicated enhanced inhibitory control with no difference between the single language L2 use and the dual language context relative to the baseline L1 game. Interestingly, retrospective ratings of effort involved indicated that single language L2 game and the dual language game were considered equally difficult as compared to the single language L1 (baseline) game. In this bilingual sample L2 may well demand increased suppression of L1 with little modulation by switching. It is noteworthy I think that the effects observed arose despite the time interval between end of a game and the initiation of the two tasks. In fact, English (the L2) was used to convey all instructions [bar those respecting the type of language game] and to address any questions. In that sense a suppressive state of L1 Polish may have been sustained. However, this would operate not only after the L2 single language context and dual language context games but also after the L1 Polish (baseline) game itself potentially reducing ERP effects and contributing to the null behavioral effects.

Overall, the study is well-motivated, the design is sound, and details of the appropriate analyses are well-reported and informatively discussed. The study notes that the outcomes are likely to be sample contingent and so offer a boundary condition on the claim that a dual language context elicits increased inhibitory control demands compared to a single language context. Most importantly the study emphasizes the

value of neural indices in advancing our understanding of language control.

The ACH is an hypothesis about the dynamics of language as modulated by the nature of the interactional context: test, refutation and extension at the end of the day requires tracking these processes in real-time. Study 3 begins to consider properties of a test that might then allow pertinent signals to be tracked in real time. Study 3 pursues this matter only indirectly with respect to questions of language control in bilingual speakers. Instead it examines the extent to which co-occurring sources of verbal conflict (i.e., within-trial sources of conflict) lead to enhancements in control revealed in the ERP trace and in manual response to colour-word Stroop conflict when this conflict is combined with so-called word-word conflict (the presence of a non-target stimulus word) in close perceptual proximity. Prior research (Rey-Mermet, Gade & Steinhauser, 2019) indicated reduced effects of conflict for co-occurring conflict (ie., within trial) that tapped different processing demands (spatial and semantic) with support for a sequential resolution of these conflicts in the ERP profile and evidence for their interaction in the behavioral data. The aim of Study 3 was to examine outcomes when the co-occurring, ie., within-trial, conflicts plausibly tap a common representational or processing substrate - a semantic component. The within-subjects design with a sample size of $n = 67$ is sound, the analysis well-conceived and the outcomes sensibly interpreted. In the event, the ERP data revealed an additive effect of the two sources of conflict on the N450 component, traditionally linked to the resolution of verbal conflict, and here indicative of more optimal conflict resolution. Interestingly, by contrast, but in line with the previous research (Rey-Mermet et al.), the behavioural data indicated an interaction between the two sources of conflict consistent with the processes of resolving Stroop conflict and word-word interference acting sequentially. Discussion rightly concerns the relative immediacy of the ERP data compared to the convolution of processes involved in a behavioral response. One relevant interpretative factor here is that resolving Stroop conflict ensures a correct manual response whereas resolving the word-word interference does not directly command the response though it may be pre-requisite for its initiation.

As noted above, Study 3 does not directly follow the exploratory trajectory of Studies 1 and 2. I turn therefore to its putative integration into the thesis offered by the opening synopsis. Study 3 indicates that distinct sources of conflict can be addressed within the same time frame and may boost the efficiency with which conflict is resolved. As argued, in the synopsis, the study is pertinent to the question of language control in bilingual speakers because multiple sources of conflict can occur (e.g., between situational cues bearing on language choice and sources with a semantic basis such as the dual activation of lexical items). The synopsis makes the plausible speculation that adaptive response in bilingual speakers occasioned by a dual language context (as contrasted with other contexts) may be selectively evident in the within-trial adaptation effect. The scope for testing this speculation requires further innovative methods to induce strong language control demands. Particularly important here are likely to be developments that allow researchers to track dynamic indices of conflict resolution in real-time with continuous rather than punctate task-demands so that research can explore the cognitive and linguistic dynamics of conversation.

Yours truly

A handwritten signature in black ink that reads "D.W. Green". The signature is written in a cursive style with a large initial 'D' and 'G'.

Dr. David W. Green, Emeritus Professor of Psychology

Experimental Psychology
Faculty of Brain Sciences
University College London
Gower Street, London WC1E 6BT
UK
Email: d.w.green@ucl.ac.uk