A Meta-Analysis of the Joint Simon Effect

April Karlinsky (april.karlinsky@alumni.ubc.ca)
School of Kinesiology, University of British Columbia
210 – 6081 University Boulevard, Vancouver, BC V6T 1Z1 Canada

Keith R. Lohse (rehabinformatics@gmail.com)
School of Kinesiology, Auburn University
301 Wire Road, Auburn, AL 36849 USA

Melanie Y. Lam (mlam@stfx.ca)
Department of Human Kinetics, St. Francis Xavier University
PO Box 5000 (Courier 1 West Street), Antigonish, NS B2G 2W5 Canada

Abstract
Since its design in 2003, the joint Simon task and corollary joint Simon effect (JSE) have been invaluable tools towards the study of joint action and the understanding of how individuals represent the action/task of a co-actor. The purpose of this meta-analysis was to systematically and quantitatively review the sizeable behavioural evidence for the JSE. Google Scholar was used to identify studies citing the first report of the joint Simon task (Sebanz, Knoblich, & Prinz, 2003) up until June 23, 2015. After screening, thirty-nine manuscripts were included in the meta-analysis, thirteen of which included individual go/no-go (IGNG) control data. Separate random-effects models were conducted for both the joint Simon and IGNG datasets, and meta-regression models were used to assess potential moderators that may impact the strength of the JSE. The results provide an important quantitative summary of the literature and serve as a foundation for future research surrounding the JSE.

Keywords: joint action; spatial compatibility; co-representation

Introduction
Throughout the day, people engage in a variety of social interactions that mold our behaviour, and even independent events can be shaped by those around us. In recent years, much research has been devoted to better understanding how individuals mentally represent the presence, tasks, and actions of others, and how such representations influence one’s own behaviour, in contrast to matched behaviours performed alone. A valuable experimental paradigm towards this end has been a spatial compatibility task, more specifically the Simon task, which can be performed in an individual (e.g., Simon, 1969) or joint setting (e.g., Sebanz, Knoblich, & Prinz, 2003).

The (Joint) Simon Task
In a typical two-choice Simon task, stimuli are presented to the left or right of centre. A non-spatial stimulus feature (e.g., colour, shape, tone pitch) informs the participant whether to make either a left or right key press response. For example, a triangle requires a left key press response while a circle requires a right key press response. Even though the stimulus location (left, right) is irrelevant to the task, it nevertheless modulates responses, such that responses are faster and more accurate when the spatial location of the stimulus and response are compatible (e.g., left-left) than when they are incompatible (e.g., left-right). This phenomenon, known as the spatial compatibility or Simon effect, has been shown to be robust, with this pattern of results replicated in many studies (for review, see Lu & Proctor, 1995).

In a social variant of the Simon task, two people are each assigned a stimulus-response mapping, such that a go/no-go protocol is completed independent from, yet complementary to the other’s task. The emergence of a spatial compatibility effect (henceforth referred to as a joint Simon effect, JSE) in the joint setting was taken as evidence that representations were formed for not only one’s own part of the task but also their co-actor’s (Sebanz et al., 2003), since the effect was noticeably absent when participants performed the same go/no-go protocol alone (individual go/no-go task, IGNG) (see Callan, Klisz, & Parsons, 1974).

Interpretations of the Joint Simon Effect
The JSE was originally interpreted in terms of the action co-representation account (e.g., Sebanz et al., 2003; for elaboration and more detailed review of this and subsequent interpretations, see Dolk et al., 2014). According to this account, individuals represent a co-actor’s task quasi-automatically; it is the representation of the alternative stimulus-response mapping that is thought to increase response conflict, eliciting the JSE. Other authors have posited the actor co-representation account, whereby response conflict emerges from the representation of the co-actor, as opposed to the co-actor’s specific task, such that conflict surrounds which agent should act when (Wanke et al., 2011). However, these accounts do not explain why the JSEs are induced in non-social contexts (e.g., Guagnano, Rusconi, & Umilta, 2010). In efforts to offer a more comprehensive explanation for the JSE, Dolk, Hommel, Prinz, and Liepert (2013) formulated and tested the referential coding account. This account posits that greater similarity across action event representations can lead to a greater emphasis on their discriminating features (e.g., location). In a series of five experiments, they manipulated the social nature of the experimental setup in two ways: (1) absence of a biological co-actor, and (2) removing any event character (e.g., sound). They showed that the JSE could be
elicited by non-social action events (e.g., Japanese waving cat) but not if the “event-like character of the sounds and movements” are eliminated (Dolk, Hommel, et al., 2013, p. 1255). What makes the referential coding account particularly appealing is that it can explain not only the occurrence of the JSE in non-social contexts, but also the more pronounced JSEs observed when there is increased self-other integration (Colzato, de Bruijn, & Hommel, 2012), as presumably within friendly partnerships (e.g., Hommel, Colzato, & van den Wildenberg, 2009), in-group interactions (e.g., Iani, Anelli, Nicoletti, Arcuri, & Rubichi, 2011; McClung, Jentsch, & Reicher, 2013; Müller, Kühn, et al., 2011), and cooperative contexts (e.g., Iani et al., 2011; Iani, Anelli, Nicoletti, & Rubichi, 2014).

Current Meta-Analytic Review

The current meta-analysis offers several novel contributions to the field of joint action. First, to our knowledge, it is the only application of quantitative methods to evaluate the substantial body of work that has emerged since the introduction of the joint Simon task (Sebanz et al., 2003). As such, it complements recent qualitative literature reviews (e.g., Dolk et al., 2014) while providing unique insights into the nature of co-representation, as indexed by the JSE. Second, we explored the size of: (1) the overall JSE, (2) the JSE when the original conditions were conceptually replicated (see Sebanz et al., 2003, Experiment 1), and (3) the JSE when an elimination or reversal of the effect was anticipated. The inclusion of these latter moderator analyses enhances our understanding of the JSE and its sensitivity to experimental manipulations. Third, we included an analysis of the IGNG task, which is considered an important control when investigating the JSE and enriches interpretations of joint effects (e.g., Sebanz et al., 2003).

Methods

Search Strategy and Study Selection

On June 23, 2015, two authors (AK and MYL) conducted an electronic search in Google Scholar for all citations of Sebanz et al. (2003). Following the addition of Sebanz et al. (2003) to the search results and removal of duplicates, 329 records were screened for eligibility. The following exclusion criteria were used to screen the articles: (a) manuscripts that were not published or translated into English; (b) manuscripts that did not include a joint Simon task; (c) studies that did not report response times (RT) and standard deviations (SD) or standard errors (SE); (d) studies examining children (<18 years old). It should be noted that articles examining joint action in special populations (e.g., individuals with schizophrenia) were not excluded, but only the data for healthy controls were included in the analyses.

Two authors (AK and MYL) screened articles by title and abstract according to these criteria. These same authors then used the criteria to screen the remaining 61 articles by full text for inclusion. When there was disagreement, the authors discussed the articles in question until consensus was reached. A total of 42 manuscripts remained eligible for inclusion in the quantitative analysis, but 3 of these manuscripts were subsequently excluded as they were doctoral dissertations whose eligible studies were also published (and included) as distinct manuscripts (Anelli, 2012; Müller, 2013; Sellaro, 2013). The 39 manuscripts remaining in the meta-analysis comprised 104 independent groups of participants (contributing 95 joint Simon datasets and 35 IGNG datasets), as some manuscripts contained multiple experiments and/or multiple groups of participants.1

Data Extraction

Two authors (AK and MYL) independently extracted data from each manuscript relevant to sample size, experimental manipulation, and response time (means and SDs or SEs).2 When necessary, data were manually estimated from reported figures. These two authors discussed any discrepancies between their extractions until consensus was reached with respect to the data included in the analyses.

Data Analysis

Cohen’s $d$ was calculated directly from the extracted RT data and the pooled between-subject SD. In cases of repeated measures designs, data were averaged across conditions such that each independent group of participants contributed only one effect size to each analysis. The effect sizes and variances were entered into a random-effects meta-analysis using the ‘metafor’ package in R (R Core Team 2014; Viechtbauer, 2010) and the DerSimonian and Laird method of estimation (Borenstein, Hedges, Higgins, & Rothstein, 2011). Effect size calculation was arranged such that effects favouring a JSE always had a positive value (i.e., incompatible mean RT - compatible mean RT). An effect size of zero indicated no difference between compatible and incompatible trials.

Custom scripts were written to test random-effects models for the overall effect of spatial compatibility within joint Simon and IGNG tasks (Cooper, Hedges, & Valentine, 2009), and Egger’s test of asymmetry was used to assess bias (Egger, Davey Smith, Schneider, & Minder, 1997). Considering the wide range of experimental manipulations within the joint Simon task literature, we also conducted two moderator analyses using meta-analytic regression. First, conditions conducted as controls (control moderator) were compared to all other conditions,3 to provide an index of the JSE unmodulated by experimental variables. Second,

---

1 In total, the data of 2079 and 583 participants went towards the joint Simon ($M = 21.88$/group, $SD = 9.02$) and IGNG ($M = 16.66$/group, $SD = 5.79$) random-effects meta-analyses, respectively.

2 When the number of participants per group was not specified, the total number of participants reported was assumed to be distributed evenly amongst groups. Standard errors (SE) were converted into standard deviations (SD) for future computations.

3 Control condition criteria included a physically present, human co-actor, actively responding to an alternative stimulus.
conditions hypothesized by the original authors to eliminate or reverse the JSE (wipeout moderator) were compared to all other effect sizes.\footnote{In cases where authors provided alternative hypotheses regarding whether the JSE would manifest itself, or not, we could not definitively code the condition as a 'wipeout' or 'non-wipeout,' and the dataset was excluded from the analysis (n = 10).} \footnote{Large samples were defined as n > 24, reflecting the 75th percentile of sample size.} Unlike the "overall" random-effects model of the JSE, in cases of repeated measures designs, we preferentially submitted a group’s ‘control’ or ‘wipeout’ data (when available) towards the relevant meta-regression model (rather than averaging across within-group conditions). Details regarding the raw data, moderator coding, and analysis scripts are available online at https://github.com/keithlohse/social_simon_meta.

Results

No Spatial Compatibility Effect in IGNG Contexts

As expected, the IGNG studies (n = 35) yielded no evidence of a spatial compatibility effect (i.e., the RT difference between incompatible and compatible trials was not statistically different from zero), d = 0.07, 95% confidence intervals (CI) [-0.01, 0.16]. A statistical test of asymmetry revealed the distribution was not skewed, t(33) = -0.76, \( p = .45 \).

Evidence of Positivity Bias and Small Effect Sizes Across Joint Simon Studies

Prior to analysis, a funnel plot revealed an extremely positive and imprecise effect size (from Dolk et al., 2012, see data point in bottom right corner of Figure 1A) which was removed from all subsequent analyses.

Figure 1A shows the distribution of joint Simon task effect sizes as a function of the standard error in each study (n = 94). Even with the Dolk et al. (2012) data point removed, a statistical test of asymmetry confirmed the positive skew in these data, t(92) = 3.25, \( p = .002 \), indicating significant bias, with more small, positive studies being published. The random-effects model summary effect size was \( d = 0.26 \), 95% CI [0.21, 0.30].\footnote{Considering the significant positive skew across the dataset, we also ran a second random-effects model restricted to large samples in efforts to remove bias.\footnote{Restricted to the largest studies (n = 20), the distribution was not skewed, t(18) = 0.96, \( p = .35 \), and the summary effect size was reduced, \( d = 0.17 \), 95% CI [0.10, 0.25].}}

No Evidence Control Conditions Moderate the JSE

We used meta-regression to compare the effect sizes derived from control conditions (n = 23) to all other conditions (n = 71), to broadly assess any modulation of the effect by experimental manipulations. There was no significant difference between the effect sizes of control conditions, \( d = 0.34 \), 95% CI [0.24, 0.44], compared to non-control conditions, \( d = 0.24 \), 95% CI [0.19, 0.29], \( p = .074 \). As shown in Figure 1B, the distribution of the effect sizes remained significantly skewed, t(92) = 3.33, \( p = .001 \).

Wipeout Conditions Decrease the JSE

Considering that ‘non-control’ conditions encompass both those experimental designs hypothesized to augment and to diminish the JSE, we conducted an additional meta-regression model to assess any moderating effects of conditions explicitly hypothesized by the original authors to eliminate or reverse the JSE (n = 16) compared to those that were not (n = 68). The summary effect size of wipeout conditions \( d = 0.12 \), 95% CI [0.01, 0.22], was significantly smaller than that of non-wipeout conditions \( d = 0.33 \), 95% CI [0.29, 0.48].\footnote{In the meta-regression of the wipeout moderator (triangles = controls; circles = non-controls), Figure 1C shows the distribution of joint Simon task effect sizes as a function of precision (standard error) for A) overall random-effects model; B) meta-regression of the control moderator (triangles = controls; circles = non-controls); and C) meta-regression of the wipeout moderator (triangles = wipeouts; circles = non-wipeouts). Positive values show a difference in favour of a JSE (i.e., faster RTs on compatible trials).}
Discussion

Since its design in 2003, many researchers have used the joint Simon task to explore the nature, extent, and boundaries of shared representations, as indexed by the JSE (for review, see Dolk et al., 2014). The present meta-analysis provides the first, much-needed quantitative summary of the literature, and serves both as a snapshot of the research to date, and a foundation on which to build future inquiries.

Across 39 manuscripts, our meta-analysis suggests the JSE is a reliable, albeit small, effect (summary $d = 0.26$). However, this analysis also revealed significant asymmetry within the data, potentially indicative of publication bias. Specifically, the data are positively skewed (even after removing an outlier), such that more small “positive” studies are being published than those with “negative” results. When we limited our analysis to large samples, the distribution was no longer skewed, but it revealed that the “real” effect size is likely even smaller than it first appeared ($d = 0.17$). This has two principle implications: (1) researchers studying this effect need an adequate sample size to achieve statistical power, and (2) there is probably limited “practical” significance of this effect, although it is still useful as a behavioural assay to understand cognitive processes (when conducted with sufficient power).

The small JSE effect size also reinforces the importance of the IGNG random-effects model, where we confirmed that a compatibility effect did not arise under individual task conditions. It should be noted that of the 39 manuscripts eligible for the joint Simon analysis, only 13 included an IGNG condition. The failure to include such a condition is of potential concern as it has been shown that a small but significant spatial compatibility effect can be observed in a go/no-go task (see Callan et al., 1974). In the case that a significant effect is found in the IGNG condition, then this compromises interpretations of the JSE.

Given the sizeable body of research included in the random-effects model of the JSE, we sought to parcel out factors that could be moderating the size of the JSE. We began with an exploration of control versus non-control task conditions. The meta-regression analysis revealed no evidence that control conditions yielded JSEs that were reliably different to those of experimental conditions. A possible explanation for our finding is that task conditions have been manipulated to elicit a range of effects on the JSE (e.g., reverse, eliminate, decrease, increase), which could result in cancellation effects and account for the lack of statistical difference between the size of the JSE under control and non-control conditions. Another plausible interpretation is that the JSE is sufficiently robust that there is some leeway in what one can do experimentally and still elicit the JSE.\footnote{\textsuperscript{7} A supplementary analysis of the control versus non-control task conditions, with the wipeout data removed from the latter to diminish potential cancellation effects, also yielded no significant difference between the summary effect sizes. This supports the notion that there is some flexibility in the task conditions that can be applied and still elicit the JSE.}

As a next step, we classified experimental conditions anticipated to eliminate or reverse the JSE as ‘wipeout’ conditions, and used a meta-regression model to assess their potential moderating effect on the size of the JSE. As anticipated, the summary effect size of the wipeout conditions was significantly smaller than the non-wipeout conditions. However, we wish to add a note of caution when interpreting this analysis. Our coding was based on the original authors’ predictions, which we assume to represent \textit{a priori} hypotheses, but it is possible some were made \textit{a posteriori}, reflecting post hoc justifications for the findings (Kerr, 1998; Lohse, Buchanan, \& Miller, 2016). An important message to convey to authors is to ensure they are transparent about whether their hypotheses are \textit{a priori} or \textit{a posteriori}. In the case that a hypothesis is generated based on theory or prior research, then they should be clear to outline why they believed a condition would eliminate or reverse the JSE. Alternatively, if after data collection potential explanations for what has been found are devised, then authors should be upfront about this. While \textit{a posteriori} hypotheses tend to be looked at less favourably, they do offer a springboard to test other methods or experimental designs. Nevertheless, the current results confirm that the JSE is sensitive to manipulations ‘designed’ to diminish its presence.

As the first quantitative description of the joint Simon literature, a clear future direction would be to meta-analytically capture the studies not included here, for example with respect to special populations (e.g., de la Asuncion, Bervoets, Morrens, Sabbe, \& De Bruijn, 2015; Liepelt et al., 2012). Additionally, and particularly in light of the asymmetry present in the current data, subsequent researchers could attempt to solicit unpublished ‘file drawer’ data, which might help to counter the observed positivity bias, and provide a more accurate picture of the conducted research and estimate of the underlying effect. Also missing from the present analyses are studies not reporting enough data to calculate their associated effect sizes (e.g., no error bars on figures, or not specifying what measure the error bars represent). As such, we urge researchers and reviewers to be diligent towards the reporting of all results, to avoid such omissions in the future.

As a final note, we encourage researchers who are designing an experiment to investigate the JSE to perform (and report) an \textit{a priori} power analysis (Cumming, 2012; Lohse et al., 2016). A shortcoming of some joint Simon studies is the inadequate sample size. Indeed, across all the

\textsuperscript{6} The summary effect size of wipeout conditions was also significantly smaller than that of control conditions, but we have omitted this additional analysis for brevity, given that the reported difference between wipeout and non-wipeout conditions is stronger evidence of the former’s impact on the size of the JSE.
studies included in this meta-analysis, not one reported estimating sample size. The effect sizes presented in this paper could be used to conduct a power analysis, and this simple procedure will help ensure that the JSE that is (or is not) being detected is a real effect. Since the joint Simon task is commonly used to explore joint action and co-representation, it is of great import to establish that the observed effect is appropriately powered if we are to infer its underlying mechanisms and influence on behaviour.

References

(* indicates manuscript included in the meta-analysis)


