Title

Assessing affective modulation of intentional binding effect: A 2AFC Psychophysics experiment with emotional words

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Toro Hernandez, Felipe D
Gabiatti, Victor Nogueira Dias
Ramalho, Theresa Moraes
et al.

Publication Date

2024

Peer reviewed

Assessing affective modulation of intentional binding effect: A 2AFC Psychophysics experiment with emotional words

Felipe Diego Toro-Hernández[1] (felipe.toro@ufabc.edu.br), Victor Nogueira Dias Gabiatti[1] (victor.gabiatti@ufabc.edu.br), Theresa Moraes Ramalho[2] (ramalho.theresa@aluno.ufabc.edu.br), Andre M. Cravo[2] (andre.cravo@ufabc.edu.br) & Peter M. E. Claessens[2] (peter.claessens@ufabc.edu.br)

[1]Graduate Program in Neuroscience and Cognition, Federal University of ABC (UFABC)
[2]Center of Mathematics, Computation and Cognition (CMCC), Federal University of ABC (UFABC)

Abstract

Intentional binding (IB) is the experience of temporal interval compression between voluntary actions and subsequent events when the latter are perceived to be caused on purpose by the agent's actions. It can be measured experimentally by comparing the judgments of temporal intervals between either a voluntary act or an external event, and a later sensory consequence. Evidence suggests this might be modulated by the emotional valence of the consequence. However, controversies have arisen over the consistency of the results and the methodology they were obtained with. Here, we aimed to measure this affective modulation using a two-interval forced-choice (2AFC) discrimination task and word stimuli. Three factors were employed: agency (agency and passive), emotional valence (neutral, positive, and negative words), and interval duration ratio determined based on individual values of just noticeable differences (JND). Participants had to judge which of two intervals presented in each trial was shorter. Generalized linear mixed model analysis indicated that there was an effect of IB, but no affective modulation. Dissociation of component mechanisms of SoA are discussed to better understand results and suggest further directions.

Keywords: Intentional binding; Sense of Agency; temporal cognition; cognitive-affective neuroscience; psychophysics; neurolinguistics.

Introduction

Sense of agency: Definition and relevance

Sense of agency (SoA) is the feeling that one's actions are related to external consequences, creating a sense of phenomenological coherence between events that are generated as consequences of the agent's own versus other peoples' actions (David, Newen & Vogeley, 2008). Even before we act, our goal-oriented motor system prepares the organism to perceive consequences in the external environment, to prepare the self for sensory events that are generated by the action of our own body (Moore, 2016).

This has been theorized to be the result of anticipation of sensory outcomes, either by a comparison between expected and actual outcome (Frith, Blakemore & Wolpert, 2000), by mental inferences of causation of authorship (Wegner, 2003), or, more recently, by a Bayesian based cue integration mechanism (Moore & Fletcher, 2012). In psychopathology, these estimations of agency fail in some syndromes, in which patients have a distorted agentic experience. We can find one example in the case of learned helplessness (Seligman, 1972), in which people underestimate their influence in generating external consequences (Moore, 2016). Another example can be found in delusions of control in schizophrenia, in which patients think that their own actions could be externally generated by other agents (Frith, 2005).

Importantly, in healthy subjects, the valence of the outcome of an action is reported to modulate agentic experience. This perceptual bias is known as "self-serving bias" (Bradley, 1978), because people tend to attribute to themselves agency more in situations that generate positive rather than negative outcomes, thus preserving self-esteem and stability to the self (Greenberg et al. 1992).

This phenomenon is particularly relevant for dissociating explicit and implicit levels of SoA (Moore, 2016; Synofzik, Vosgerau & Newen, 2008): we may evaluate the consequences of our actions based on processes involving more high-order reflective interpretation of the involvement of ourselves in external consequences (explicit SoA), or just "feel" that we are agents of contingencies in our environment, involving low-level sensorimotor processing (implicit SoA) (Moore & Obhi, 2012).

While self-serving bias may be expressed in explicit evaluations of agency (e.g., thinking about the outcome of an action performed days ago), implicit measures may shed light on how this bias works on a pre-reflective perceptual level. Following this thought, Haggard, Clark and Kalogeras (2002) found that, when subjects had to judge the moment of either a voluntary action or an outcome (agency condition), or two baseline conditions – a voluntary action without outcome, and a sensory event without a preceding voluntary action – participants experienced a subjective temporal compression between voluntary action and outcome. This effect was called intentional binding (IB). To quantify it, Haggard et al. (2002) use a paradigm based on Libet's clock (Libet et al. 1983). In this procedure, observers monitor a rotating clock hand. They report the estimated instant of either the voluntary action or the external consequence, in both agency and baseline conditions, by means of the position of the pointer on the clock face (Haggard et al. 2002). The presence and size of the IB effect is inferred from a forward shift of the estimate of the instant of voluntary action towards the moment of the external consequence, and a backward shift of the estimated moment of a consequence towards the voluntary action. According to Haggard et al. (2002), this does not happen

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society*.

when the 'consequence' event is unrelated to the agent's actions.

Although this effect has been proposed as a measure of implicit SoA and has been replicated in a similar way several times (Haggard et al. 2002; Moore & Obhi, 2012; Yoshie & Haggard, 2013; Yoshie & Haggard, 2017), it is not exempt from criticism. Evidence suggests that it may also happen in the absence of voluntary action, notably in the context of causal expectations ("causal binding" effect, see Buehner, 2012; Buehner & Humphreys, 2009). Nonetheless, temporal compression between an interval and a subsequent contingency seems to be stronger in the presence of voluntary action (Hoerl et al. 2020), with evidence suggesting that this effect may need both agency and causal attributions to emerge (Cravo, Claessens, & Baldo, 2009), leading to different conceptualizations of the phenomenon (Hoerl et al. 2020).

Affective modulation of intentional binding

Experimental evidence with implicit measures of agency suggests a diminished effect of intentional binding when a negative outcome is presented, in comparison with a positive outcome. The results of the study by Yoshie et al. (2013) show that, in conditions of predictability (i.e., participants were able to predict the emotional valence of the outcome), when the outcome is a negatively valued sound (e.g., the sound of someone screaming in fear), IB is reduced in comparison with positive sound stimuli, such as a vocal expression of amusement. This is suggestive from a moral reasoning point of view: because of self-serving bias, people usually tend to attribute to themselves agency in an outcome more often when it is positive than when it is negative.

Although this affective modulation of binding has been replicated in Yoshie et al. (2017), who also manipulated predictability of occurrence and valence of the outcome stimulus, and, without the predictability manipulation, Takahata et al. (2012), there are contradictory results. First, Moreton, Callan & Hughes (2017) not only tried to replicate the original results but also added other outcome stimuli, such as faces and emoticons representing emotions, but found no statistically significant differences in binding between emotional conditions. Second, considering valence outcome, predictability, and occurrence, Christensen et al. (2016) found stronger IB when neutral outcome was expected but did not occur. Third, evaluating action selection by giving participants the possibility to choose between several keyboards to perform a voluntary action, Tanaka and Kawabata (2021) found that temporal compression was stronger with negative than with positive outcomes when unpredictable. Finally, evaluating the effect of intended outcome in binding, with participants themselves choosing the emotional valence of outcomes, Sarma and Srinivasan (2021) also showed that the intentional binding effect was larger for negative than for positive facial expressions, when they were both intended.

The problem of methodology

In summary, literature suggests that the IB effect might be modulated by the emotional valence of outcomes. However, given that the results are inconclusive, considering that different research groups document different directions of the effect, the exact role of valence is still uncertain. More data need to be collected in replications and variations of the original experiments to shed more light on whether and under which conditions this modulation occurs. Furthermore, all these studies use the Libet clock or interval estimation procedures (Moore and Obhi, 2012), that are arguably rather prone to participant bias.

While the first method consists of measuring time by means of a rotating clock, the second one consists of participants reporting estimates of their perceived interval time duration. For example, in Moreton et al. (2017), participants had to scroll a slider in a range of time estimates from 0-1000 ms to report the perceived duration of interval between action and consequence. Similarly, Sarma et al. (2021) used a numeric keyboard in which participants had to choose between 1 and 9, displayed as a multiple of hundreds on the screen (i.e., number 1 was represented as 100ms).

Notwithstanding the replicability of the results obtained with these procedures based on verbal report, the fact that all require conscious evaluation might induce a subject bias in results, which might be avoided by using psychophysical procedures that are probably less transparent to the volunteering observers (see Cravo et al. 2009; Cravo, Claessens, & Baldo, 2011; Fereday, Buehner, & Rushton, 2019; Nolden, Haering, & Kiesel, 2012). Because duration judgements are more prone to distortion unrelated to the variables of interest, Fereday et al. (2019) proposed an experimental method based on the discrimination of the time intervals in two successive presentations in causal and non-causal conditions, based on a 2AFC (two-alternative forced choice) procedure, and found effects of subjective temporal compression, which these authors attribute to impaired temporal resolution in a causal inference context.

Finally, all these studies used sound, faces, and emoticons (representing faces) to capture the emotional valence of presented outcomes, which limits its results to a specific kind of emotional representation and a limited variability in content (e.g., Yoshie et al. 2013, 2017, used only 8 sounds to represent emotions). A larger range of stimuli might be useful to better clarify controversial results and to capture more variability in emotional processing. Specifically, linguistic stimuli are known to be processed early in the brain (Pulvermüller, Shtyrov, & Hauk 2009) and have well-detailed emotional characterization by psycholinguistic norms in different languages (de Oliveira, Janczura, & Castilho, 2013; Speed & Brysbaert, 2023; Warriner, Kuperman & Brysbaert, 2013; Stadthagen-Gonzalez et al. 2017). Evidence shows that emotional words elicit ERPs (event-related potentials) that are processed early (as P1 and N170), suggesting emotional processing that may be like that of faces (Zhang et al. 2014). In fact, emotional valence of words may be dissociated of semantics, as our brain is more prompt to recognize stimuli that are potentially dangerous for the organism (Zhang et al. 2014), in line with results that suggest unconscious processing of the former to be more automatic than the latter (Lei et al. 2017). Thus, studies assessing affective modulation of emotional valence may benefit from the variability provided by word stimuli by using several words with similar valence values.

The present study intends to clarify controversial results by adapting the original experiment performed by Yoshie et al. (2013), using a 2AFC paradigm and adding a new stimulus type (linguistic emotional stimuli) to better understand the affective modulation of the intentional binding effect, avoiding the mentioned possible methodological challenges. To our knowledge, this is the first time that the intentional binding is measured using word stimuli.

Methods

A 2AFC psychophysics procedure was designed with word stimuli. First, 180 word stimuli were selected from Brazilian Portuguese databases (de Oliveira, et al. 2013), to obtain psycholinguistic variables of interest (length and normative data on concreteness and frequency) and create emotional valence groups (positive, negative, and neutral). Second, an online, pre-experimental questionnaire was designed to capture familiarity, and age of acquisition, variables that were not included in word databases. Third, another pre-experimental questionnaire was designed to capture emotional valence and arousal judgments, in order to confirm the characterization of emotional categories. Finally, a 2AFC experiment was performed in the psychophysics laboratory at Universidade Federal do ABC. The whole procedure received approval from the local Research Ethics Committee (CAAE: 62401922.5.0000.5594). Participants did not receive any kind of incentive and they all gave written informed consent. The data are available at: https://osf.io/93ya6/.

Psycholinguistic variables: Analysis and questionnaires

Selection of linguistic stimuli Psycholinguistic parameters (emotional valence, arousal, concreteness, frequency, and word length) of 908 words were obtained from Brazilian Portuguese databases (Estivalet & Meunier, 2017; Janczura et al. 2007; de Oliveira et al. 2013). Because low-frequency words have different response times according to emotional valence (Barriga-Paulino et al. 2022), and to avoid confounding effects of different content words, such as differences between processing verbs or nouns, only nouns and high-frequency words, defined as words with more than 40 occurrences per million, were selected.

The final stimulus selection was composed of 180 words. Words with the highest values of emotional valence (based in a rating scale between 1-9; de Oliveira et al. 2013) were selected as positive (e.g., "happiness"), and words with the lowest values were selected as negative (e.g., "cocaine"). To select neutral stimuli, the mean value of valence was calculated (5.46), and the words centered on this central value were selected as neutral (e.g., "object"). Emotional word categories were matched in concreteness, frequency, and word length, so they only varied in emotional valence and arousal.

Emotional valence and Arousal analyses Statistics were performed using the aov function in R (v. 4.1.1, R Core Team, 2023). Repeated measures ANOVA analyses were performed. The variable "emotional valence" was created including the three emotional categories, positive, negative, and neutral. These categories were well differentiated, showing a main effect of emotional valence ($F(2,177) = 5.092$, $p < .001$, $\eta 2 = 0.98$). Bonferroni post-hoc comparisons confirmed that the positive group of words ($M = 7.99$, $SD = 0.29$) had higher ratings of valence than neutral ($M = 5.45$, $SD = 0.24$), and negative words ($M = 2.16$, $SD = 0.40$). Furthermore, neutral words ($M = 5.45$, $SD = 0.24$) had higher ratings of valence than negative words ($M = 2.16$, $SD = 0.40$). Emotional valence categories also had a main effect of arousal ($F(2,177) = 490.2$, $p < .001$, $\eta 2 = 0.85$). Bonferroni HSD post-hoc comparisons showed that negative words ($M = 7.28$, $SD = 0.66$) were more arousing than neutral ($M = 4.6$ 9, $SD = 0.59$), and positive words ($M = 3.32$, $SD = 0.84$). Also, neutral words ($M = 4.69$, $SD = 0.59$) were more arousing than positive words ($M = 3.32$, $SD = 0.84$). There were no differences between emotional valence categories for concreteness, frequency, and word length.

Online questionnaire 1

Participants and procedure Twenty-three native Brazilian Portuguese speakers (13 female, 10 male, $M_{age} = 26.7$, $SD_{age} = 8.10$) answered questions about familiarity and age of acquisition of the target words. To capture Familiarity and Age of acquisition, a Google Forms questionnaire was constructed involving the 180 words selected for this study. Participants were required to answer how familiar they were with each target word on a scale from 1-7 (Gernsbacher, 1984), and the age at which they thought they had learned the word using one of the reference intervals (0-2 years old, 3-4, 5-6, 7-8, 9-10, 11-12, 13+) (Bird, Franklin & Howard, 2001). Age of acquisition ratings were also transformed into a 1-7 scale (Bird et al. 2001).

Results of questionnaire 1 Statistics were performed using the aov function in R (v. 4.1.1). Repeated measures ANOVA analyses were performed. Familiarity ratings showed a main effect of emotional valence ($F(2,177) = 13.3$, $p < .001$, $\eta 2 = 0.13$). Bonferroni post-hoc analyses showed that positive words ($M = 5.99$, $SD = 0.64$) were more familiar than neutral words ($M = 5.64$, $SD = 0.87$), and more familiar than negative words ($M = 5.27$, $SD = 0.80$). Also, neutral words ($M = 5.64$, $SD = 0.87$) were more familiar than negative words ($M = 5.27$, $SD = 0.80$). Furthermore, Age of acquisition ratings also presented a main effect of emotional valence ($F(2,177) = 5.319$, $p < .01$, $\eta 2 = 0.06$). Bonferroni post-hoc analyses showed that positive words (M = 2.92, SD = 1.10) were reported to be acquired earlier in life than neutral (M = 3.48, SD = 1.19), and negative words (M = 3.51, SD = 1.05).

Online questionnaire 2

Participants and procedure Thirty-one native Brazilian Portuguese speakers (19 female, 12 male, M_{age} = 29.4, SD_{age} = 7.4) answered questions about emotional valence and arousal of the target words. A Google Forms questionnaire was constructed involving the 180 words selected for this study. Participants were required to answer how pleasant (from 1-9: very unpleasant to very pleasant) and how arousing (from 1-9: very relaxing to very arousing) the set of words was (de Oliveira et al., 2013).

Results of questionnaire 2 Statistics were performed using the aov function in R (v. 4.1.1). Repeated measures ANOVA analyses were performed. Emotional valence showed a main effect of emotional valence (F(2,177) = 1.249, p <.001, $\eta 2$ = 0.93). Bonferroni post-hoc comparisons confirmed that the positive group of words (M = 7.96, SD = 0.41) had higher ratings of valence than neutral (M = 5.72, SD = 0.61), and negative words (M = 2.44, SD = 0.76). Furthermore, neutral words had higher ratings of valence than negative words. Emotional valence categories also produced a main effect on arousal ratings (F(2,177) = 50.15, p <.001, $\eta 2$ = 0.36). Bonferroni HSD post-hoc comparisons showed that negative words (M = 6.03, SD = 0.48) were more arousing than neutral (M = 5.15, SD = 0.46), and positive words (M = 4.42, SD = 1.37). Also, neutral words (M = 5.15, SD = 0.46) were more arousing than positive words (M = 4.42, SD = 1.37).

Conclusions of the survey results

Emotional valence was well characterized in the positive, negative, and neutral conditions. The results with Questionnaire 2 confirm that the selected word stimuli vary in emotional categories. Analyses of psycholinguistic variables yielded no significant differences between categories, except for arousal, that was highly correlated with emotional valence in the original data of de Oliveira et al. (2013), leading to arousal differences across emotional categories (also showed in data obtained from questionnaire 2). This makes it difficult to manipulate emotional valence independently of arousal.

Furthermore, data collected by means of questionnaire 1 showed a difference across conditions for both familiarity and age of acquisition. This difference should be considered in posterior analyses performed evaluating the affective modulation of the IB effect. If the effect is found (i.e., different binding effects regarding emotional valence of outcomes), the complementarity of these variables may confound the results.

Main experiment

Participants, apparatus, and stimuli Thirty-two participants (15 female, 17 male, M_{age} = 24, SD_{age} = 4.57) were enrolled in this experiment, based on the sample size of similar studies in which this effect was found (Yoshie et al. 2013, 2017). It was programmed with the OpenSesame platform (Mathôt, Schreij & Theeuwes, 2012) and was carried out in a dark, quiet sound-proof room on a Samsung CRT computer monitor (SyncMaster 997mb), set to a vertical refresh rate of
160Hz and a resolution of 640 x 480 pixels, localized at approximately 40 cm from participants' faces. To counterbalance for the effect of stimulus length, nonwords were generated as a sequence of random string characters (e.g., “democracy” might be matched to the string “lxeikias”).

Design and procedure In this experiment, we aimed to capture the participant’s proportion of “shorter” responses through a perceptual discrimination 2AFC procedure. Three factors were manipulated: Condition (agency and passive), duration of the second interval (5 durations: either -2, -1, 0, 1, or 2 fractional JNDs of the first interval), and emotional valence (positive, negative, and neutral). Each block had a total of 90 trials, giving a total of 360. Trials could be of agency or passive, and each of them had two intervals to be compared. In agency trials, participants had to press a mouse

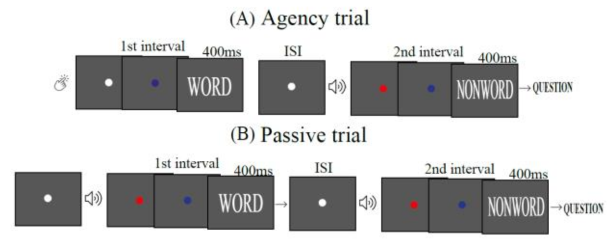

Figure 1: Depiction of trials for both conditions

button to generate a word after a duration as the first interval. Then, the presentation of an inter-stimulus-interval (ISI; with the apparition of a white dot for 500 ms) preceded the second interval, in which participants heard a sound, followed by the presentation of a nonword (Figure 1A). In passive trials, participants did not press any button but instead heard a sound before both stimulus presentations, separated by the ISI as in agency trials. The first intervals always produced words and the second intervals always presented nonwords in both conditions (Figure 1B). Words and nonwords were presented for 400 ms. After each pair of intervals, volunteers had to answer which of both intervals was shorter.

Finally, the words used were presented in sequences of equal emotional valence. Thus, a sequence of 30 positive, another of 30 neutral, and one of 30 negative words were shown in each block, twice, thus giving a total of four blocks (with 90 words each). Emotional stimuli were disposed this way to avoid confounding effects of emotion presentation or fatigue, and to reproduce the predictability of valence that produce under which valence modulation was originally found (Yoshie et al. 2013). These four blocks, two agency and two passive conditions, were permutated randomly.

While the first interval was sampled from a uniform random distribution between 200 and 300ms, the duration of the second interval was calculated as a fraction of the first, in multiples of JND estimates based on a 3-up-1 down and 3-down-1-up adaptive staircase procedure. This procedure was performed with only passive trials at the beginning of the experiment, with a length of 20 trials in each staircase. The results were fitted with a logistic psychometric function with a guess and lapse rate both fixed at 1%, in a procedure implemented in Python. From the model fits, the just noticeable difference (JND) between two intervals was defined as half the difference between the interval ratios that produce 25% and 75% responses of the first interval being shorter/longer. Thus, the second interval could be either -2, -1, 0, 1, or 2 fractional JNDs of first interval.

Results

Data Analysis

A generalized linear mixed-model (GLMM) analysis was performed using the glmer function of the lme4 package in R (Bates et al. 2015), with proportion of "first interval shortest/second interval longest" responses and interval duration ratio, expressed in JNDs, as fixed factors, and participants as a random factor for intercept and main effects. The GLMM with a logistic link function is equivalent to a mixed-model approach to logistic psychometric function fitting, with guess and lapse rates set to 0. The fixed-effects slope of the psychometric function was significantly positive ($\beta = 0.991 \pm .11$, $z = 9.14$, $p < .001$), indicating that volunteers gave increasing numbers of responses indicating the first interval as the shortest with increasing ratios of the second over the first interval. In other words, participants were able to perform the task, and the adaptive procedure was successful in calibrating difficulty across volunteers (Figure 2). For the same interval ratios, a main effect of agency condition was found, since the first interval was judged as shorter more often in agency trials ($\beta = 0.139 \pm .07$, $z = 1.98$, $p = .048$), indicating underestimation of self-generated intervals, consistent with intentional binding as a shift to the left of the psychometric curve (Figure 2). Furthermore, an interaction model between agency and emotional valence yielded no significant results, neither for negative vs neutral words ($\beta = 0.03 \pm .10$, $z = 0.29$, $p = .769$), nor for positive vs neutral words ($\beta < 0.001, \pm .10$, $z = 0.001$, $p = .999$). Finally, there was no global effect of emotional valence on the estimation of the intervals, as the comparison between the GLMM models with and without emotional valence yielded no significant results (likelihood ratio test, $\chi 2(2) = 3.95$, $p = .86$) (Figure 3). Thus, the results did not suggest the expected affective modulation of the intentional binding effect.

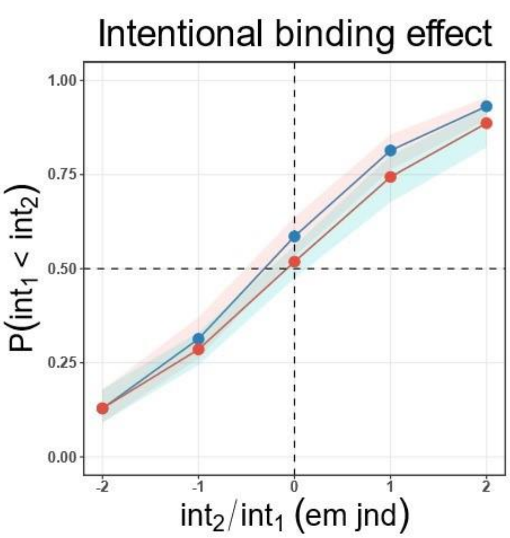

Figure 2: Psychometric curves of main effect of IB. In agency condition a = agency (higher line), p = passive.

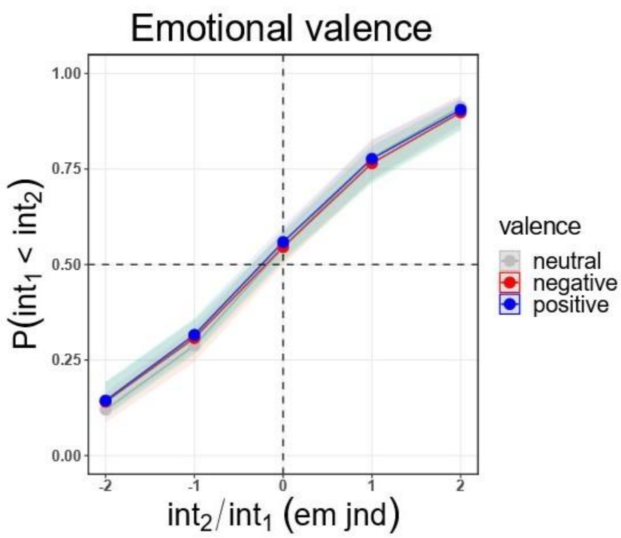

Figure 3: Psychometric curves of main effect emotional valence. Notice that there are no differences.

Conclusions

In this study, we aimed to look for the affective modulation of IB, i.e., different magnitudes of binding depending on valence, with a 2AFC procedure using linguistic stimuli as outcome. Results show that participants properly understood the task, and data showed a tendency to underestimate agency intervals in comparison with passive intervals of the same duration, thus replicating IB effect (Haggard et al. 2002). However, consistent with the results of Moreton et al. (2017), we did not find any significant effect of valence, meaning that emotional categories (positive, neutral, or negative) did not influence participants' interval estimation.

General discussion

Affective modulation of IB was described by Yoshie et al. (2013), who wanted to explore self-serving bias effects in implicit SoA. They hypothesized that implicit measures might be important to explore the relation between agency and emotional outcomes, and found stronger IB effects for positive and neutral, than for negative outcomes. In the same vein, Takahata et al. (2012) designed a procedure in which sound stimuli were associated with monetary losses (negative), no losses at all (neutral), or monetary wins (positive). In this study, they also found that negative outcomes had an attenuating effect on IB in comparison with positive and neutral outcomes.

However, the exact mechanism by which this modulation might occur is debatable, since SoA implies several mechanisms which may interfere (Haggard, 2017; Moore, 2016; Wen & Imamizu, 2022). One way in which sense of agency may arise is by active inference (prospective account), implying that the brain captures predictions about certain outcomes possibilities related to our voluntary actions or bodily movements (Moore & Fletcher, 2012).

Following this rationale, Yoshie et al. (2017) claim that active inferences created by the voluntary motor system may cause the affective modulation of IB, since actions would be directed to generate positive outcomes. Thus, they compared conditions in which participants were either able or unable to predict the emotional outcome of an action, using the Libet clock to capture IB, finding stronger binding for positive than for negative outcomes only when predictable. When the valence of the outcome was unpredictable, they did not find any differences, pointing at the role of prospective processing in the affective modulation of IB (Yoshie et al. 2017).

Interestingly, studying the dissociation between prospective and retrospective components of SoA, Christensen et al. (2016) found bigger IB for positive outcomes in retrospective-unpredictable conditions (i.e., participants received a sound after a voluntary action in 75% of intervals), deepening phenomenon understanding.

On the one hand, the emotional direction of the retrospective component was in the same line as the quoted experiments (Takahata et al. 2012; Yoshie et al. 2013; Yoshie et al. 2017). However, this did not exactly match with the prospective interpretation given by Yoshie et al. (2017). On the other hand, the results of Takahata et al. (2012) were obtained in conditions of unpredictability, thus raising the question of what the actual contribution of predictability and outcome valence in the modulation of IB is.

Furthermore, Tanaka et al. (2021) found a stronger IB effect for negative than for positive outcomes only when stimuli were unpredictable, yielding null results when predictable. In one condition participants had the possibility of choosing one key to perform voluntary action, and in other they only could press several. Because results were obtained only when choosing among several keys, action selection showed to have an influence on the affective modulation of IB. In the same line, by manipulating intention (i.e., participants could choose what emotion they wanted to see as an outcome), Sarma et al. (2021), also found stronger IB for negative than for positive emotional outcomes when they were intended. This shows how the interaction of different mechanisms of SoA interact with emotional valence, in a way to anticipate external responses, either for self-preservation or self-serving bias.

One can think of two ways of interpreting the absence of evidence of any affective modulation in our own study. First, following the results of Takahata et al. (2012), possibly word stimuli were not stimulating enough to generate the expected effect. However, this would not explain that Moreton et al. (2017) did not find the same effect as Yoshie et al. (2013), neither with emoticons, faces nor sounds. In that case, as Moreton et al. (2017) stated, the effect could be small or unreliable. Second, following the line of Tanaka et al. (2019) and Sarma et al. (2021), our study did not involve the control of action selection, or intention of outcome, which may be nullifying the modulation of emotional outcomes.

As affective modulation may reflect a self-serving bias, one may think that watching words on a screen would lead to a weaker effect than that of sounds or faces, because people would not associate the meaning of words with their voluntary actions. Nevertheless, there is evidence suggesting that emotional words might be processed in a similar fashion to that of faces (Zhang et al., 2014). In fact, negative words may have a strong impact to prepare organism to harmful interactions with the environment (Zhao et al., 2018).

Our methodology presents some important strengths to highlight. First, the 2AFC experiment implemented here can be used as a proxy for more "objective" interval comparisons, since all studies quoted before used subjectively biased methodologies, such as the Libet clock and interval estimation approaches (Moore & Obhi, 2012; see also Fereday et al. 2019). Second, the incorporation of word stimuli could be used as a new way of measuring emotional valence in IB, since it allows to obtain more variability in measures, for example when several words are used rather than a small number of emotional sounds.

Our results shed light onto the understanding of attributional bias in the processing of affective modulation of IB. The absence of evidence for an affective modulation suggests that implicit agency evaluations may not be driven by emotion, meaning that we would need explicit high-order evaluations of consequences for the existence of self-serving bias (i.e., to be more prompt to accept as ours results that are positive, in comparison with negative). Processing time compression between voluntary action and outcome on the one hand, and emotion on the other, may be dissociated, thus needing a posteriori integration.

This work raises important questions about linguistic stimuli in the processing of agency and the prediction of emotional valence in IB. Further directions may incorporate action selection and intention of outcome, to understand how these influence valence effects in IB. Also, neuroimaging and psychophysiology studies are needed to dissociate the mechanisms underlying SoA and its relationship with IB.

Acknowledgments

We appreciate the help of all professors and colleague at the Timing UFABC lab, whose comments were of fundamental in the development of this project. This work was supported by a UFABC scholarship and funding by CAPES to the Posgraduate Program in Neuroscience and Cognition at UFABC.

References

Barriga-Paulino, C. I., Guerreiro, M., Faísca, L., & Reis, A. (2022). Does emotional valence modulate word recognition? A behavioral study manipulating frequency and arousal. *Acta psychologica*, 223. doi: 10.1016/j.actpsy.2021.103484.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bird, H., Franklin, S. & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers* 33, 73–79. doi: 10.3758/BF03195349.

Bradley, G. (1978). Self-serving biases in the attribution process: A reexamination of the fact or fiction question. *Journal of Personality and Social Psychology*, *36*(1), 56–71. doi: 10.1037/0022-3514.36.1.56.

Buehner, M. J. (2012). Understanding the past, predicting the future: Causation, not intentional action, is the root of temporal binding. *Psychological science*, 23(12), 1490-1497. doi: 10.1177/0956797612444612

Buehner, M. J., & Humphreys, G. R. (2009). Causal Binding of Actions to Their Effects. *Psychological Science*, 20(10), 1221-1228. doi: 10.1111/j.14679280.2009.02435.x

Christensen, J. F., Yoshie, M., Di Costa, S., & Haggard, P. (2016). Emotional valence, sense of agency and responsibility: A study using intentional binding. 39 *Consciousness and Cognition*, 43, 1-10. doi: 10.1016/j.concog.2016.02.016.

Cravo, A. M., Claessens, P. M. E., & Baldo, M. V. C. (2009). Voluntary action and causality in temporal binding. *Experimental Brain Research*, 199(1), 95–99. doi: 10.1007/s00221-009-1969-0.

Cravo, A. M., Claessens, P. M., & Baldo, M. V. (2011). The relation between action, predictability and temporal contiguity in temporal binding. *Acta Psychologica*, 136(1), 157-166. doi: 10.1016/j.actpsy.2010.11.005.

David, N., Newen, A., & Vogeley, K. (2008). The "sense of agency" and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, 17(2), 523–534. doi: 10.1016/j.concog.2008.03.004.

Estivalet, G. L., & Meunier, F. (2017). Corpus psicolinguístico Léxico do Português Brasileiro. *Soletras*, 33, 212-229. doi: 10.12957/soletras.2017.29702.

Fereday, R., Buehner, M. J., & Rushton, S. K. (2019). The role of time perception in temporal binding: Impaired temporal resolution in causal sequences. *Cognition*, 193, 104005. doi: 10.1016/j.cognition.2019.06.017.

Frith, C. (2005). The self in action: Lessons from delusions of control. *Consciousness and cognition*, 14(4), 752-770. doi: 10.1016/j.concog.2005.04.002.

Frith CD, Blakemore SJ, Wolpert DM. (2000). Abnormalities in the awareness and control of action. *Philos Trans R Soc Lond B Biol Sci*. 29;355(1404), 1771-88. doi: 10.1098/rstb.2000.0734.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. Journal of *Experimental Psychology: General*, 113(2), 256–281. doi: 10.1037/0096-3445.113.2.256.

Greenberg, J., Pyszczynski, T., Burling, J., and Tibbs, K. (1992). Depression, self-focused attention, and the self-serving attributional bias. *Pers. Individ. Dif.* 13, 959–965. doi: 10.1016/0191-8869(92)90129-D.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 196-207. Haggard, P. (2017). Sense of agency in the human brain. Nature Reviews Neuroscience, 18(4), 196-207. doi: 10.1038/nrn.2017.14.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382-385. doi: 10.1038/nn827.

Hoerl, C., Lorimer, S., McCormack, T., Lagnado, D. A., Blakey, E., Tecwyn, E. C., & Buehner, M. J. (2020). Temporal binding, causation, and agency: Developing a new theoretical framework. *Cognitive Science*, 44(5), e12843. doi: 10.1111/cogs.12843.

Janczura, G., Castilho, G., Oliveira, N., van Erven, T., & Huang, T. (2007). Normas de concretude para 909 palavras da língua portuguesa. *Psicologia: Teoria e Pesquisa*, 23, 195-204. doi: 10.1590/S0102-37722007000200010.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain*, 106, 623–642. doi: 10.1093/brain/106.3.623.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324. doi:10.3758/s13428-011-0168-7.

Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, 7, 1272. doi: 10.3389/fpsyg.2016.01272.

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Conscious Cogn*. 21(1), 59-68. doi: 10.1016/j.concog.2011.08.010.

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and cognition*, 21(1), 546-561. doi: 10.1016/j.concog.2011.12.002.

Moreton J., Callan M., & Hughes G. (2017). How much does emotional valence of action outcomes affect temporal binding? *Consciousness and Cognition*, 49, 25–34. doi: 10.1016/j.concog.2016.12.008.

Nolden, S., Haering, C., & Kiesel, A. (2012). Assessing intentional binding with the method of constant stimuli.

Consciousness and Cognition, 21(3), 1176–1185. doi: 10.1016/j.concog.2012.05.003.
de Oliveira, N. R. D., Janczura, G. A., & Castilho, G. M. D. (2013). Normas de alerta e valência para 908 palavras da Língua Portuguesa. Psicologia: Teoria e Pesquisa, 29(2), 185-200. doi: 10.1590/S0102-37722013000200008.
Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, 110, 81–94. doi: 10.1016/j.bandl.2008.12.001.
R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Sarma, D., & Srinivasan, N. (2021). Intended emotions influence intentional binding with emotional faces: Larger binding for intended negative emotions. *Consciousness and Cognition*, 92, 103136. doi: 10.1016/j.concog.2021.103136.
Seligman, M. E. (1972). Learned helplessness. *Annual review of medicine*, 23(1), 407-412.
Speed, L.J., Brysbaert, M. (2023). Ratings of valence, arousal, happiness, anger, fear, sadness, disgust, and surprise for 24,000 Dutch words. *Behav Res*. doi: 10.3758/s13428-023-02239-6.
Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M.A. et al. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behav Res* 49, 111–123. doi: 10.3758/s13428-015-0700-2.
Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and cognition*, 17(1), 219-239. doi: 10.1016/j.concog.2007.03.010.
Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's not my fault: postdictive modulation of intentional binding by monetary gains and losses. *PloS one*, 7(12), e53421. doi: 10.1371/journal.pone.0053421.
Tanaka, T., & Kawabata, H. (2021). Sense of agency is modulated by interactions between action choice, outcome valence, and predictability. *Current Psychology*, 40(4), 1795-1806. doi: 10.1007/s12144-018-0121-3.
Warriner, A.B., Kuperman, V. & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res* 45, 1191–1207. doi: 10.3758/s13428-012-0314-x.
Wegner, D. (2003). The mind's best trick: how we experience conscious will. *Trends In Cognitive Science*, 7, 6569. doi: 10.1016/s1364-6613(03)00002-0.
Wen, W., Imamizu, H. (2022). The sense of agency in perception, behaviour and human machine interactions. *Nat Rev Psychol* 1, 211–222. doi: 10.1038/s44159-022-00030-6.
Lei Y, Dou H, Liu Q, Zhang W, Zhang Z and Li H. (2017). Automatic Processing of Emotional Words in the Absence of Awareness: The Critical Role of P2. *Front. Psychol.* 8:592. doi: 10.3389/fpsyg.2017.00592
Yoshie, M., & Haggard, P. (2013). Negative emotional outcomes attenuate sense of agency over voluntary actions. *Current Biology*, 23(20), 2028-2032. doi: 10.1016/j.cub.2013.08.034.
Yoshie, M., & Haggard, P. (2017). Effects of emotional valence on sense of agency require a predictive model. *Nature Scientific Report*, 7(8733), 1-8. doi: 10.1038/s41598-017-08803-3.
Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., ... & Luo, Y. J. (2014). Three stages of emotional word processing: an ERP study with rapid serial visual presentation. Social cognitive and affective neuroscience, 9(12), 1897-1903. doi: 10.1093/scan/nst188.
Zhao W, Chen L, Zhou C, Luo W. (2018). Neural Correlates of Emotion Processing in Word Detection Task. *Front Psychol.* 9:832. doi: 10.3389/fpsyg.2018.00832.