Breaking Control - How confusion of control can help understand the Sense of Agency

Streiling, K. (kai.streiling@tu-darmstadt.de)[1], Penaz, V.[1] & van Dam, L.C.J.[1, 2]
[1]Centre for Cognitive Science, Technical University of Darmstadt
[2]Center for Mind, Brain and Behavior, University of Marburg and Justus Liebig University Giessen

Abstract

The Sense of Agency (SoA) is a crucial component of robust human motor control and learning. To study the SoA, existing work has largely focussed on retrospective ratings about the level of perceived control over a single visible object. Here we propose a new design, using queries during a trial, for assessing participants' attribution of control to one of multiple acting objects performing the same obstacle avoidance task. Correlation analyses between the controlled object and the object to which Agency was (wrongly) attributed showed that Agency confusion is largely driven by similarities in the dynamics of participants' control input and the confounding object's behaviour. Visual attention was also more attracted to the object believed to be under the participants' control. These results are consistent with the comparator and cue-integration models of Agency and the presented approach opens up novel avenues to explore the dynamic processes underlying the SoA.

Keywords: Sense of Agency; continuous control; confusion

Introduction

Human motor control is extremely robust. We can compensate perturbations, adapt to visuomotor delay (Rohde et al., 2014) and learn novel action-effect mappings (Cunningham & Welch, 1994). An important tool that enables this robustness is the Sense of Agency (SoA) (David et al., 2008; Gallagher, 2000; Pacherie, 2007; Synofzik et al., 2008), i.e. an intuitive registration that it was us who triggered a certain event or caused a movement of an object. This causal attribution is necessary to distinguish action-related feedback, useful for learning, from externally generated stimuli. The so-called comparator model of SoA (Feinberg, 1978; Frith et al., 2000) offers a simple mechanistic explanation how such a registration is formed: each motor action is accompanied with an efference copy (von Holst & Mittelstaedt, 1950) that is used in an internal forward model (Wolpert et al., 1995) to predict the sensory consequences of the action. If this prediction is "similar" to the actual feedback, a Sense of Agency is formed. Furthermore, cue-integration approaches suggest that such prediction-outcome congruencies may be integrated with other contextual and cognitive cues to generate an SoA Moore and Fletcher, 2012; Synofzik et al., 2008.

Most experimental work has approached the SoA using discrete actions and events, such as button presses and flashes, which complicates generalisation to more natural continuous control. Paradigms exploring SoA in continuous motor control are rarer and mostly use retrospective judgements about the perceived control. Existing work had participants rate their SoA in motor control tasks whilst being presented with either their own or somebody else's feedback through a mirror or a computer display (Asai, 2015; Miyawaki & Morioka, 2020; Nielsen, 1963; Wegner et al., 2004). Moreover, participants rated their perceived control (e.g. yes/no/how much) about a single object (arms, hands, dots) that they had more or less control over. The absence of alternative objects that Agency might be attributed to might push the beliefs about Agency towards higher ratings. Some studies have approached this problem by introducing distinct alternatives (Wang et al., 2020), others also offered a behavioural correlate of the SoA in the form of tracking error (self-controlled, larger SoA, smaller tracking error). Yet the respective judgements were always collected at the end of a trial which distances them from the actual experience or makes it necessary to reduce trial duration to a few seconds.

Here, we present a novel paradigm with two important changes: participants select one of multiple acting objects they believe to control, which can reveal confusions in the SoA, and they do so repeatedly during a trial. Leaning on racing games, participants controlled one of three cars on a computer display to avoid obstacles. We used collisions of any car with an obstacle to trigger a query what car a participant thought they controlled. We can use misidentifications to explore what drove these confusions in the SoA.

In the following we show that confusion about the object under control (i) is driven by the dynamic alignment between objects as suggested by the comparator model, (ii) is reflected in participants' gaze behaviour and (iii) binds actions to fit the (wrong) belief of control.

Methods

Motivated by anecdotal evidence from multiplayer video games, where players lost track of the character they were controlling, we developed a novel continuous control paradigm to study confusion of the Sense of Agency. The task was to steer a car left and right to avoid obstacles moving downwards on a screen while other cars, called Non Player Characters (NPCs), performed the same task (see Figure 1). Furthermore, using a game controller in this task decoupled proprioceptive from visual feedback and made it easier to apply less familiar action-effect mappings, i.e. controlling a car's acceleration.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

Participants Fifteen participants with normal or corrected to normal vision took part. Two participants were excluded due to technical issues during the experiment, leaving $N = 13$ participants (7 male, 6 female, avg. age 23.4 years, 19 to 34 years). Participants were recruited from the student population and received course credits for their participation. The study was approved by the Ethics Committee of the Technical University of Darmstadt and participants provided written informed consent before participation.

Apparatus The experiment ran on a conventional desktop computer (AMD Ryzen 7 3700X 8-Core CPU and NVIDIA GeForce RTX 3060 GPU) with a Gigabyte M32U screen (resolution set to 3840×2160 px / 67.7×39.2 cm, 144 Hz) and was programmed using python 3.10 and the library pyglet 2.0. Participants were sitting at a table holding an XBox game controller in their hands and with their head on the chin rest of an EyeLink 1000 Plus' tower mount with a viewing distance of approx. 57 cm to the screen.

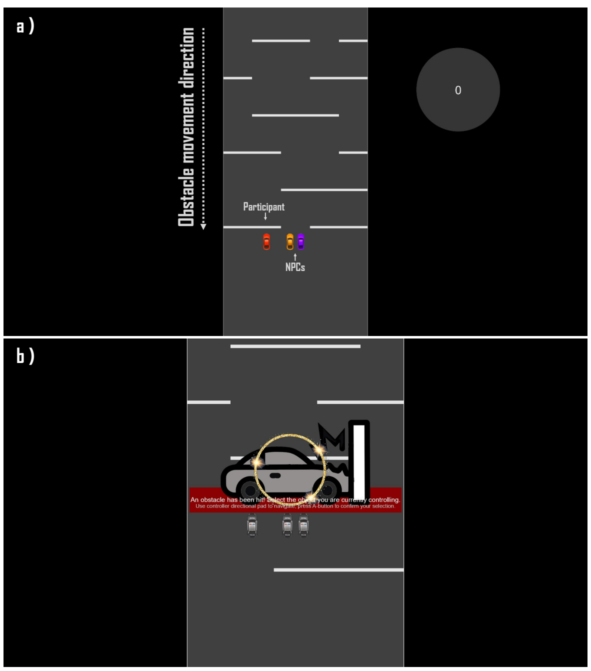

Figure 1: Task display. a) Participants were moving one of three cars left and right to avoid obstacles that moved towards them. They were awarded points for each obstacle passed. Points were displayed in the top right. b) Upon a collision of any car, the trial paused and participants had to select which object they thought they were controlling.

Stimuli and Procedure A trial scene showed a black background with a centred, grey road (approx. 17 cm wide) and a circular score area in the top right corner (see Figure 1a). The cars were placed at approx. 11.3 cm from the bottom of the screen and were approx. 0.8 cm wide, 1.9 cm high and either red, orange or purple. The obstacles were presented as white lines moving downwards (making the cars appear to move upwards). For placing the gaps in these obstacles the road was segregated into five invisible lanes. An obstacle consisted of one part (white bar, approx. 0.35 cm high) per lane and left at least one lane open as a gap. The distance between obstacles and their vertical downward speed varied with control mode and difficulty (see below). For acceleration control, they moved at approx. 4.6 cm/s and were approx. 6.5 cm apart in *hard* trials and moved at approx. 4.2 cm/s with a distance of 7 cm in *easy* trials. Sets of 5 obstacles were handcrafted and combined into different maps, such that each trial showed a different combination of obstacles.

In separate experimental blocks, participants could either control one car's acceleration, velocity or position. At the start of the experiment participants practised all three control modes in a training block to get used to the task. They trained for one trial with 50 obstacles for each control mode without additional NPCs. Afterwards, 3 measurement blocks (one per control mode) of 18 trials with 70 obstacles each were performed. Trials and blocks were presented in a random order per participant. Within each block, we varied the presence of NPCs (none or two), the difficulty of the obstacle avoidance task (*easy* or *hard*) and whether control was switched between objects during a trial with two NPCs present (*switch* or *no-switch*) yielding 6 different trial types which were repeated 3 times adding up to 18 trials per block. The EyeLink was calibrated using a 9-point display before the training and again before each block. Between blocks participants could take breaks and move freely.

Because of systematic differences between the control modes and the blocked, independent design, we only present results from acceleration controlled trials. Acceleration control was least directly influencing the cars positions, i.e. it prevented "wiggling" a car left and right to identify it, whilst still being intuitive. Therefore, and for the purpose of conciseness, we will only focus on the methods and results of the acceleration block in this manuscript. Participants could apply a horizontal acceleration of approx. 8.8 cm/s^2 in an all-or-nothing principle using the game controller's left joystick.

Non-Participant Characters (NPCs) were controlled by an adapted Linear Quadratic Regulator (LQR) which included the current position and velocity of the NPC and the position of the next two obstacles as state variables and applied a horizontal acceleration as control signal. The planning horizon T was set to include the next two obstacles and the cost function minimized the distance of the controlled object to a target position in a gap in the first obstacle at time $T/2$ and to a target position in the second obstacle at time T. The linear control law L was calculated following the implementation of (Straub & Rothkopf, 2022). To account for the vertical elongation of the cars, the cost function was extended to include the distance to the first target not only at $T/2$ but between the time when the top of the car reached the obstacle and when it had passed the obstacle by three quarters. The target positions were updated at each obstacle to generate smooth behaviour. If an obstacle had multiple gaps, the gap closest to the current

NPC position was chosen. Within a targeted gap, the concrete target position for the LQR was set to the gap's centre plus a slight random offset to the left or right drawn from a uniform distribution between 0.01 cm and 0.2 cm. NPC 1 was offset to the left and NPC 2 to the right avoiding complete overlapping of the cars. As the LQR generates optimal behaviour which might be easily distinguished from participants' own non-optimal behaviour, we introduced two parameters to decrease its performance: A *randomness* probability with which not the closest but a random gap was chosen when selecting the next target (NPC 1: *hard*: 0.11, *easy*: 0.1; NPC 2: *hard*: 0.09, *easy*: 0.12) and a *reaction time* value r that prolonged the focus of the next gap (NPC 1: *hard*: 0.01, *easy*: 0.15; NPC 2: *hard*: 0.04, *easy*: 0.04). Specifically, at each gap a factor was drawn from a normal distribution with mean r and standard deviation 0.1, clipped at 0 and multiplied with the number of time steps necessary to pass the height of the car given the current obstacle speed. The focus on the next gap was delayed by the resulting number of time steps. These parameters also allowed to create slightly differing behaviours for NPC 1 and NPC 2. Additionally, the control signal on acceleration was discretised in an all-or-nothing principle equivalent to participants' control. These manipulations led to on average 1.6 collisions per trial of NPC 1 and 2.4 collisions of NPC 2 (for comparison: participants collided on average 9.3 times).

A collision of any car with an obstacle paused the trial and started a control query (see Figure 1b). All cars were masked in grey and the colliding obstacle disappeared. Participants were tasked to "Select the car [they] are currently controlling" using the arrow keys of the game controller. After they confirmed their choice with a button press, a short countdown of 2 seconds appeared after which the trial continued. On the left and right side of the road, cars could bump into the crash barriers which would stop their horizontal movement without triggering a collision. Additionally to collision-triggered queries, control was pseudo-randomly queried up to three times per trial after 20 %, 50 % and/ or 75 % of the obstacles were passed. Such a random query was only started if the preceding two obstacles were successfully passed.

At the beginning of a trial participants were controlling the red, NPC 1 the blue and NPC 2 the orange car. In switch trials, control switched to one of the other two cars after a random query twice. The car previously controlled by the participant was then taken over by the respective LQR controller of NPC 1 or NPC 2. For example, a participant could first control the red car for the first 20 % of the obstacles of a trial, then, after a random query, the blue car until the next random query after 50 % of the obstacles after which they controlled the yellow car for the remainder of the trial. However, if a participant was repeatedly colliding no random query and thus no control switch occurred.

Data and Analysis Positions of the cars and visible obstacles were collected with on average 143 Hz when no NPCs were present. When NPCs were present the frame rate decreased to approx. 100 Hz due to the computational demands of the LQR controllers. To calculate the cars' smoothed velocity v_k at a time step k from the recorded horizontal positions x, we used Equation 1 (Engbert & Kliegl, 2003). Acceleration was calculated equivalently from the resulting velocities.

$$v_k = \frac{x_{k+2} + x_{k+1} - x_{k-1} - x_{k-2}}{6\Delta t} \tag{1}$$

Note that the time Δt between two time points k and $k + 1$ varied slightly due to the commercial monitor, which was taken into account in the implementation.

Furthermore, participants' answers to control queries were flagged as *robustly* confused, if they misidentified the controlled car and the confounding and actually controlled car were not visually overlapping, i.e. if their centres were at least 0.8 cm apart. For further analysis only such robust confusions were used. After the experiment, participants were debriefed and asked to report their applied strategies if any.

Gaze position and pupil size were stored at the respective behavioural data frequency. For both eyes, we first calculated a smoothed velocity according to Equation 1 and acceleration as the simple pointwise derivative of velocity. Samples were flagged as saccades, if for both eyes the positional difference to the previous sample was bigger than 1 degree of visual angle (deg) (Tokuda et al., 2011), eye velocity exceeded 30 deg/s (de Haas et al., 2019), or the acceleration was above 9500 deg/s^2. Samples were marked as blinks or missing, if at least one eye was gazing off the screen. Blinks were extended by leading and trailing saccades. Amongst the three cars, we estimated participants' current gaze target in a two step procedure. First, cars were marked as gaze target candidates, if they fell within an area of 0.5 deg around the current gaze position accounting for the average calibration error. Additionally, due to the obstacles coming from above, participants frequently focussed their gaze slightly above a car. We thus elongated the candidate area by the height of a car in the vertical direction. Second, from amongst the (potentially multiple) resulting candidates we estimated the gaze target to be the object with the highest positional correlation with the dominant eye over the preceding 0.5 s window. For this calculation, blinks were linearly interpolated to avoid artefacts. This procedure suggested that participants were fixating their own car in 21.6 % and one of the NPCs in 10 % of the time. All other samples were marked as having no gaze target amongst the three cars.

Furthermore, to be able to link confusion to a possible similarity between the separate cars' dynamics, we calculated the correlations between the cars' acceleration per trial over a 0.5 s right anchored sliding window. These correlations were then aggregated relative to the time points of collisions/ random queries and separately so for the switch/ no-switch conditions. For the correlation analyses we focused on the acceleration of gaze and cars, as this was the control signal applied by participants. Note however that using position or velocity for these analyses led to similar results.

To underline the results presented in the following section, we report uncorrected p values calculated using Wilcoxon signed rank tests if not stated otherwise.

Results

We first analysed the proportion of confusions (i.e. participants wrongly identified an NPC as the car they were controlling) in no-switch and switch trials. The switch trials, in which the participant's control could switch to a different car, were introduced to create a higher potential for confusion. This manipulation also allowed us to assure the validity of participants' reported control. Figure 2 shows that the number of collisions increased in switch trials, i.e. performance decreased, ($W = 0.0$, $p = 0.0002$, Fig. 2a) and also the proportion of confused reports significantly increased ($W = 4$, $p = 0.0017$, Fig. 2b). Thus, participants' reports about the controlled car were indeed connected to the actual control they could execute and vice versa also reflect its misattributions.

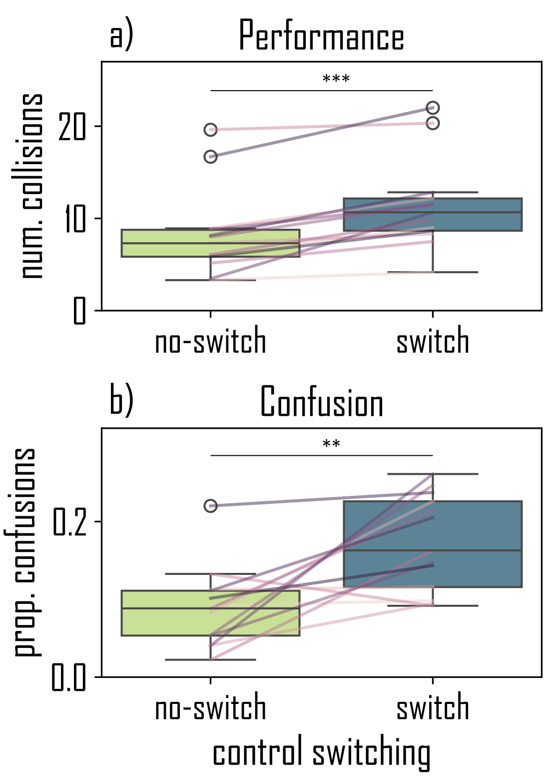

Figure 2: Performance and confusion results. a) shows the number of collisions in no-switch and switch trials (average per participant). A higher number of collisions means worse performance. b) displays the proportion of confused reports in no-switch and switch trials (average per participant). The boxes show the median, 25 % and 75 % quartiles; whiskers include datapoints within 1.5 times inter-quartile range across participants. Lines indicate individual participant averages. Stars indicate significant differences of $p < 0.01$ (2 stars) and $p < 0.001$ (3 stars).

We furthermore investigated what drove the misattributions of control, i.e. confusions in the SoA, by exploring participant and NPC behaviour before collisions. Figure 3A shows the acceleration correlations between cars in a one second time window before collisions split by participants' answers in the corresponding query. In no-switch trials (Fig. 3A top row) we find a trend that acceleration correlations between the participant's and confounding car were higher (i.e. more similar dynamics) shortly before the collision compared to the other car. These differences were however not significant in no-switch trials when using a Wilcoxon signed rank test on the correlation averages over a 0.5 s window preceding the collision (Confusion with NPC 1 (middle panel): *Part.-NPC 1* vs. *Part.-NPC 2* (purple vs. orange line) $W = 14$, $p = 0.3594$; confusion with NPC 2 (right panel): $W = 1$, $p = 0.123$). For switch trials the differences become stronger and do partially reach significance (Fig. 3A bottom row; confusion with NPC 1: *Part.-NPC 1* vs. *Part.-NPC 2* (purple vs. orange line) $W = 9$, $p = 0.0645$; confusion with NPC2: $W = 8$, $p = 0.0061$).

Figure 3B displays the proportions that each car was estimated to be the gaze target in the time leading up to a collision. Here, a trend can be observed that participants focussed their gaze more on the object they believed to control than on the others. Before collisions where control was correctly attributed, participants' visual attention was clearly on the controlled object (left panel). If, however, there was a confusion, the preceding gaze distribution is less apparent. In no-switch trials (top row), there is no distinguishable gaze target before confused collisions. Yet, in switch trials (bottom row) there was a trend of the confounding car being focussed on more than the other two, particularly when the confusion occurred with NPC 1 (comparing 0.25 s window aggregates before collisions: *NPC 1* vs. *Part.*: $W = 0$, $p = 0.0039$ and *NPC 1* vs. *NPC 2*: $W = 5$, $p = 0.0781$) but not with NPC 2 (*NPC 2* vs. *Part.*: $W = 12$, $p = 0.1309$ and *NPC 2* vs. *NPC 1*: $W = 9$, $p = 0.1289$).

Lastly, we focused on the queries after which the control did actually switch to a different car. Figure 4 shows the inter-car acceleration correlations when control switched to NPC 1 (Fig. 4b) and NPC 2 (Fig. 4c). For comparison Figure 4a shows consistent correlations before and after a query, if control did not switch to a different car afterwards. However, after a control switch to NPC 1 (Fig. 4b), there is an increased correlation of the participant's acceleration (which was previously NPC 1) and NPC 1's acceleration (which was previously controlled by the participant but now by the LQR associated with NPC 1). This difference between correlations just before and after a query is significant when control switched to NPC 1 ($W = 0$, $p = 0.0078$). For switches to NPC 2 there is a similar trend which was not significant ($W = 6$, $p = 0.1094$).

Discussion

In this study we propose confusion between multiple acting objects as a novel way to study the Sense of Agency. We present a new experimental paradigm in which participants steered a car (game-item) through an obstacle field in the presence of NPC cars and reported which item they felt they

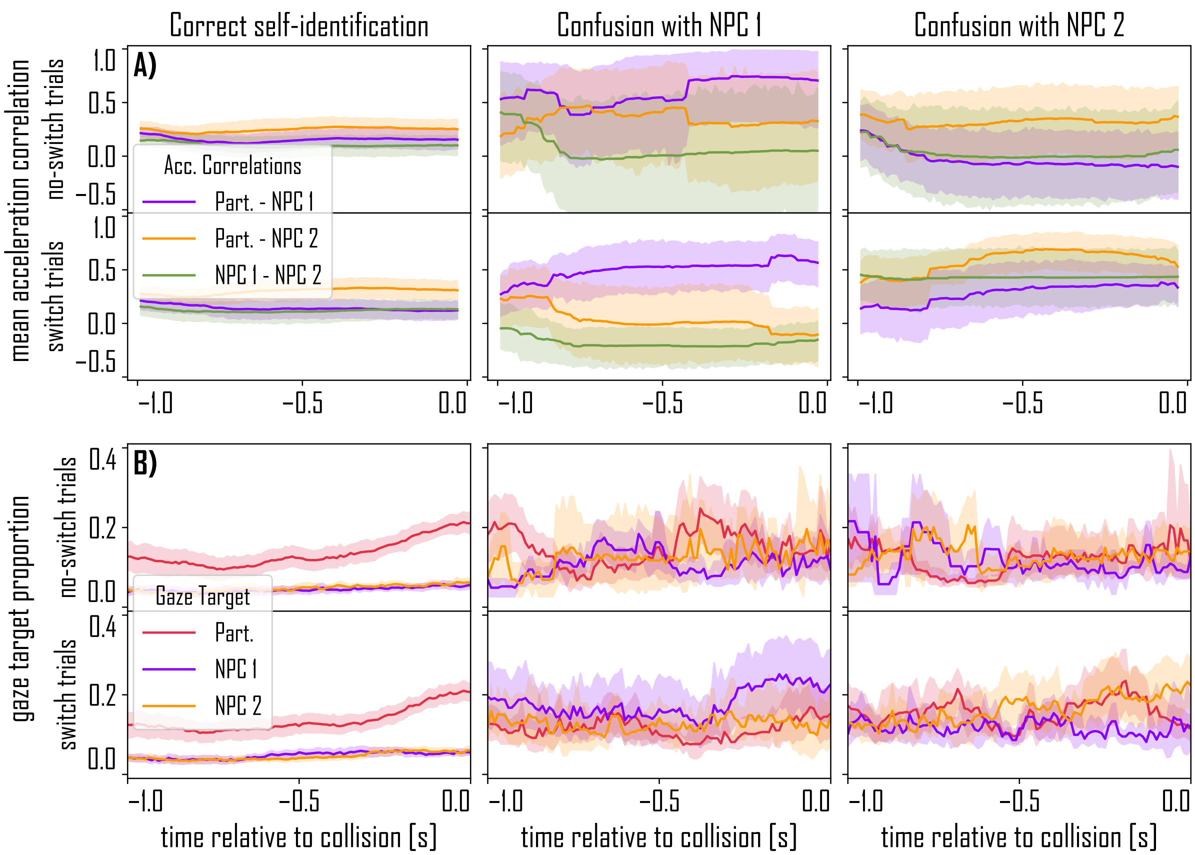

Figure 3: Similarity in car and gaze dynamics relative to collisions. A) shows the correlation of two cars' acceleration over a sliding window of 0.5 seconds for no-switch trials (top row) and switch trials (bottom row). Lines mark the mean acceleration correlations of participant and NPC 1 (purple), participant and NPC 2 (orange) and NPC 1 and NPC 2 (green). B) shows the proportion that a car was the estimated gaze target before a collision for no-switch (top row) and switch trials (bottom row). From left to right, A) and B) display correlations for the 1 second before a collision (at time 0) where the controlled object was correctly identified, confused with NPC 1 or NPC 2. Lines show the mean across participants and shaded areas mark the 94 % confidence interval calculated using bootstrapping as implemented in (Waskom, 2021).

were controlling upon a collision of any car. This way multiple SoA reports were collected during a continuous visuomotor control task. By switching the controlled car in some trials we created a higher potential for confusion and made participants challenge their otherwise strong belief of controlling the red car, with which they started each trial. Using this approach we found support for the idea that confusion is driven by the alignment in dynamic behaviour between the cars in the display. This confusion was to some degree also reflected in participants' gaze behaviour. Conversely, falsely believing a car to be under the participant's control led to behaviour consistent with this confused object.

Our results support the idea of a predictive process in forming an SoA as suggested by the comparator model: For a single controlled object, action-effect predictions and observations must be similar to form an SoA. For multiple objects we should attribute our SoA to the one with the most similar behaviour to our predictions. A confusion between objects is more likely, if (i) the prediction noise increases or (ii) dynamic behaviour of the other objects become more similar to the controlled one. While we did not quantify participants' prediction noise, we ensured sufficient uncertainty by using a game controller instead of, for instance, a graphics tablet. Using acceleration control through the game controller decoupled proprioceptive from visual feedback and created a less familiar action-effect mapping, rendering predictions more uncertain. Moreover, the increased correlation between the confounding and the actually controlled car before a confusion (Fig. 3A) provides evidence that dynamic alignment is a driving factor of SoA.

However, such a comparison, between predicted feedback and actual result, offers likely just one of multiple Agency cues that are integrated to form an SoA (Moore & Fletcher, 2012; Synofzik et al., 2008). These additional cues could explain the differences we find between confusions with NPC 1 and NPC 2. For example, in no-switch trials participants controlled the red, NPC 1 the purple and NPC 2 the orange car. We could speculate that because purple and red are less similar than orange and red, there is a lower default potential for confusion

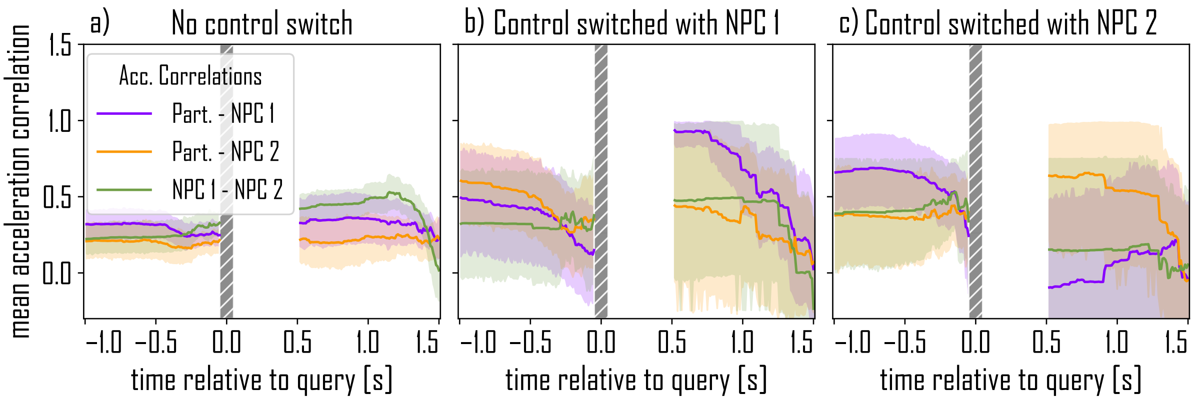

Figure 4: Similarity in car dynamics relative to random queries. Acceleration correlations are shown for random queries (from both switch and no-switch trials) after which the control did not switch (a), switched to NPC 1 (b) or NPC 2 (c). The vertical grey bar indicates an arbitrary time between the start of the query and the trial restart after participants gave their report. Because correlations are calculated over of a 0.5 s sliding window, no data is displayed in the 0.5 s after the query ended. Lines show the mean across participants and shaded areas mark the 94 % confidence interval calculated using bootstrapping (Waskom, 2021).

with NPC 1 based on its colour. This would then require a stronger dynamic alignment, i.e. a stronger correlation, in order for SoA to switch to NPC 1 and explain the trend towards a higher correlation with NPC 1 before confusion with it occurs (Fig. 3A, top row, middle panel) than with NPC 2 (Fig. 3A, top row, right panel). Such a process would be in line with Bayesian cue integration models weighting information sources according to their uncertainty (Miyawaki & Morioka, 2020; Moore & Fletcher, 2012).

Our analysis of eye-tracking data revealed that visual attention is indicative of perceived control and potential confusions thereof. Sampling information about the object one believes to control is sensible to refine estimates about its current position and velocity. Other cars however were also visually inspected and more so during times of confusion and higher uncertainty about the controlled object, i.e. before a confused collision report (see Fig. 3B). One explanation might be that participants compared their action-effect predictions with each car to update their belief about their controlled car.After a control switch, participants were more consistently misattributing their control to an NPC which is reflected in their increased visual focus on the confounding car (see Fig. 3B, bottom row).

These beliefs about control, that reflect the SoA, may be formed based on participants' actions, but they also guide action selection. These interactions might occur on different time intervals. Figure 3A shows that confusion can be driven by similarities in dynamics shortly before the reported confusion. Figure 4b and c on the other side show that a strong belief, e.g. "I am controlling the red car", persists even after control switched to another car. As a result, actions might initially be taken based on this belief, which might explain the increased correlations after a control switch between the previously controlled car's and switched-to car's acceleration. This correlation might persist until at some point a discrepancy in the predictive action-outcome comparison becomes apparent and the belief is updated. This suggests a slower or on-demand updating of the beliefs of the object under control, e.g. once predictions and effects become too dissimilar.

Querying SoA during a continuous control trial offers new ways to analyse dynamic influences on the formation of SoA. However, we also face new challenges when designing such studies: First, there are *retroactive effects* in the participants' responses to consider. Some participants reported in the debriefing that it was only at the moment of collision and when the query started that they realized which object they were controlling. Thus, a query itself can break an SoA confusion. Measuring a behavioural result of a confusion (e.g. a prolonged realization of a collision) instead of querying it, could mitigate this problem. Second, there is a choice regarding the query *timing*. We chose collisions as query time points, arguing that confusion of control makes a collision more likely. However, querying only when the participant collided would add another source of identity information and querying at every collision overly extended the experiment duration and might have also fractured the experience of *continuous* control. Thus, a more selective indicator of potential confusion to trigger an explicit query is needed. On a general note, future work could test the generalisability of the presented results by further increasing the sample size.

Taken together the presented approach shows promise for investigating the SoA in continuous control. Particularly, focusing on confusions with other game objects solving the same task in the same display opens up new avenues to investigate the role of beliefs, appearance and dynamical cues to Agency attribution.

Acknowledgments This work was supported by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) under Germany's Excellence Strategy (EXC 3066/1 "The Adaptive Mind", Project No. 533717223).

References

Asai, T. (2015). Feedback control of one's own action: Self-other sensory attribution in motor control. *Consciousness and Cognition*, *38*, 118–129. https://doi.org/10.1016/j.concog.2015.11.002

Cunningham, H. A., & Welch, R. B. (1994). Multiple concurrent visual-motor mappings: Implications for models of adaptation. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 987–999. https://doi.org/10.1037/0096-1523.20.5.987

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. https://doi.org/10.1016/j.concog.2008.03.004

de Haas, B., Iakovidis, A. L., Schwarzkopf, D. S., & Gegenfurtner, K. R. (2019). Individual differences in visual salience vary along semantic dimensions [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *116*(24), 11687–11692. https://doi.org/10.1073/pnas.1820553116

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035–1045. https://doi.org/10.1016/S0042-6989(03)00084-1

Feinberg, I. (1978). Efference Copy and Corollary Discharge: Implications for Thinking and Its Disorders [Place: US Publisher: National Institute of Mental Health]. *Schizophrenia Bulletin*, *4*(4), 636–640. https://doi.org/10.1093/schbul/4.4.636

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action [Publisher: Royal Society]. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1404), 1771–1788. https://doi.org/10.1098/rstb.2000.0734

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21. https://doi.org/10.1016/S1364-6613(99)01417-5

Kahl, S., Wiese, S., Russwinkel, N., & Kopp, S. (2022). Towards autonomous artificial agents with an active self: Modeling sense of control in situated action. *Cognitive Systems Research*, *72*, 50–62. https://doi.org/10.1016/j.cogsys.2021.11.005

Miyawaki, Y., & Morioka, S. (2020). Confusion within feedback control between cognitive and sensorimotor agency cues in self-other attribution. *Attention, Perception, & Psychophysics*, *82*(8), 3957–3972. https://doi.org/10.3758/s13414-020-02129-5

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59–68. https://doi.org/10.1016/j.concog.2011.08.010

Nielsen, T. I. (1963). Volition: A New Experimental Approach. *Scandinavian Journal of Psychology*, *4*(1), 225–230. https://doi.org/10.1111/j.1467-9450.1963.tb01326.x

Pacherie, E. (2007). The Sense of Control and the Sense of Agency. *Psyche*, *13*(1), 1. https://hal.science/ijn_00352565

Rohde, M., van Dam, L. C. J., & Ernst, M. O. (2014). Predictability is necessary for closed-loop visual feedback delay adaptation. *Journal of Vision*, *14*(3), 4. https://doi.org/10.1167/14.3.4

Straub, D., & Rothkopf, C. A. (2022). Putting perception into action with inverse optimal control for continuous psychophysics (J. Diedrichsen, J. I. Gold, & J. Diedrichsen, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, *11*, e76635. https://doi.org/10.7554/eLife.76635

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. https://doi.org/10.1016/j.concog.2007.03.010

Tokuda, S., Obinata, G., Palmer, E., & Chaparro, A. (2011). Estimation of mental workload using saccadic eye movements in a free-viewing task [ISSN: 1558-4615]. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4523–4529. https://doi.org/10.1109/IEMBS.2011.6091121

van Dam, L. C. J., & van Ee, R. (2005). The role of (micro)saccades and blinks in perceptual bi-stability from slant rivalry. *Vision Research*, *45*(18), 2417–2435. https://doi.org/10.1016/j.visres.2005.03.013

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, *37*(20), 464–476. https://doi.org/10.1007/BF00622503

Wang, S., Rajananda, S., Lau, H., & Knotts, J. D. (2020). New measures of agency from an adaptive sensorimotor task. *PLOS ONE*, *15*(12), e0244113. https://doi.org/10.1371/journal.pone.0244113

Waskom, M. L. (2021). Seaborn: Statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious Agency: Experiencing Control Over the Movements of Others. *Journal of Personality and Social Psychology*, *86*(6), 838–848. https://doi.org/10.1037/0022-3514.86.6.838

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An Internal Model for Sensorimotor Integration [Publisher: American Association for the Advancement of Science]. *Science*, *269*(5232), 1880–1882. https://doi.org/10.1126/science.7569931