Degree project

UNIVERSITY OF SKÖVDE

Sense of Agency and Automation: A Systematic Review

Master Degree Project in Cognitive Neuroscience

Second Cycle 30 credits

Spring/Summer Term 2023

Student: Ismael Albutihe

Supervisor: Andreas Kalckert

Examiner: Pilleriin Sikka

Abstract

Technological advancements have resulted in highly automated systems that are featured in many kinds of tools and devices, such as self-driving cars, autopilot in airplanes, and much more. Such systems have enabled tools to plan, decide, and act autonomously. This breakthrough resulted in a new manner of interacting with tools, known as "Human-Robot Joint Action" or "human-AI interaction," in which people and automated tools share control over the tasks that must be performed. However, little is known about the impact of such interactions on people and their sense of agency (SoA) as well as how much autonomy to grant to tools. As a result, the objective of this systematic review is to investigate and understand how automated tools affect human SoA, and if tools with different levels of automation affect our SoA differently. A search in two databases, Scopus, and MEDLINE EBSCO was conducted, and 8 articles were included. The findings suggest that the more automated the tool is, the less SoA participants experience, and that varied levels of automation may impact human SoA depending on the nature of the task. However, this topic is still in its infancy and more research is needed.

Keywords: Automation, Sense of Agency, joint action.

Sense of Agency and Automation: A Systematic Review

Trying to distinguish whether sensory events are generated by us or the environment is a critical feature that humans need to navigate and survive in this world. For example, it is important to know that it is you who is driving the car and not the passenger next to you. This feature is called Sense of Agency (SoA). SoA refers to the subjective experience of directing/steering one's actions and, as a result, external events (Gallagher, 2000). SoA is considered one of the most fundamental and significant components of human phenomenology, and in the last couple of decades, it has been extensively studied. According to research in this field, our SoA can be affected by various factors, such as human-human joint action. The term "Human-Human Joint Action" refers to when two or more individuals coordinate their actions to complete a task collectively, which typically requires precise temporal and spatial cooperation (Goswami & Vadakkepat, 2019). In such circumstances, it becomes difficult for the actor and co-actor to determine on a pre-reflective level who is accountable for the action that they have been performing (Obhi & Hall, 2011; Wegner & Wheatley, 1999). For this and other reasons, cognitive science and engineering have become more interested in joint action and SoA.

The field of engineering is continually evolving, resulting in new technologies, such as Artificial Intelligence Systems (AIS), remote-control devices, and automated tools, e.g., self-driving cars and robots, with which humans have begun to interact with on a daily basis (Wen & Haggard, 2018). The primary intent of such tools is to operate alongside users to assist them in lessening their cognitive and physical load while also boosting their performance, and these tools usually require human-machine collaboration. This has led to a new type of joint action that is sometimes called “Human-Robot Joint Action” and some other times it is called “Human–AI Interaction” (Wen & Imamizu, 2022). The main purpose of this systematic review is to investigate if the SoA could be affected by the new innovations (such as automated tools and robots) that require people to coordinate their actions around them to achieve a desirable goal.

What is the Sense of Agency?

Haggard and Chambon (2012) defined "sense of agency" as the subjective experience of controlling one's own actions and, through them, events in the external environment. According to the authors, individuals typically have a feeling of being in control of their actions most of the time, which they refer to as the normal SoA. Most of our actions are associated with distinct subjective sensations that differ in content and value. These may include the experiences of planning to act, choosing one action over another, and beginning or initiating the action. These sorts of experiences are primarily cognitive in nature and have been related to frontal lobe action planning and a motor command from the primary motor cortex and they are categorized as 'central' experiences (Passingham & Wise, 2014). Further, the SoA is often coupled with a different type of experience that is related to the body actually moving and is communicated through the activation of peripheral somatosensory receptors. Remarkably, involuntary motions usually result in these 'peripheral' feelings but not central experiences (Haggard, 2017).

The SoA could influence our daily life in different ways. For example, the intentional binding (IB) effect demonstrates how the SoA may alter our experience of time. This effect was first described in 2002, and it refers to a reduction in the apparent time delay between an intentional activity and its consequences when they are produced by oneself (Haggard et al., 2002). Sensory attenuation is another well-known effect of the SoA that illustrates how self-generated tactile stimuli are experienced to be less intense than tactile stimuli caused by others (Blakemore et al., 1998). Visual attention could also be influenced by the SoA, in a situation where control of items and occurrences is uncertain, attention is naturally drawn first to things that are under control (Wen et al., 2019; Wen, Shibata, et al., 2020; Wen, Shimazaki, et al., 2020). Furthermore, SoA could influence our daily life on a behavioral level. Such an effect could be noticed in our action selection, for example, ten-week-old newborns can detect the relationship between their leg movements and the motions of a mobile connected to their leg and respond by increasing their leg movements (Rovee & Rovee, 1969).

Theories to Explain Sense of Agency

SoA is a crucial aspect of human experience, as it refers to the feeling of control over one's actions and their consequences. Despite its significance, the exact process through which this phenomenon emerges remains uncertain. Researchers have proposed three main theories to explain the cognitive and neural underpinnings of agency, each with distinct emphases and approaches. In order to better understand the development of these theories and their contributions, it is helpful to examine each theory in more detail.

The first theory, the comparator model (Frith, 2005; Frith et al., 2000), posits that the SoA arises from the brain's ability to generate sensory predictions based on motor commands and compare these predictions with actual sensory feedback. In this model, the brain creates an efference copy of a motor instruction, which serves as a sensory prediction. This prediction is then compared to the sensory feedback received from the sensory system and the external environment. When the prediction and sensory feedback align, SoA is experienced. Conversely, when there are significant discrepancies between the prediction and sensory feedback, SoA loss occurs, often referred to as prediction error or sensorimotor incongruence (Blakemore et al., 1998, 2002; Frith et al., 2000). This model highlights the importance of the brain's ability to predict the consequences of actions and monitor discrepancies between predicted and actual outcomes. However, it does not address the temporal aspects of the SoA or consider how multiple sources of information are integrated.

The second theory, the retrospective theory (Wegner et al., 2004; Wegner & Wheatley, 1999), shifts the focus to the temporal aspects of the SoA. According to this theory, not all signals within the sensorimotor system that contribute to the SoA become available simultaneously. Some sensorimotor signals, such as those associated with action selection, are generated before an action occurs (premotor signals). Others may be generated after an action occurs due to delays in the causal chain between action and outcome or delays in receiving and processing reafferent sensory information. Consequently, the retrospective theory posits that the SoA arises from cognitive inferences based on the integration of these temporally separated signals, suggesting that one's ideas and thoughts produce one's actions (Wegner, 2003; Wegner et al., 2004). While this theory accounts for the temporal aspect of the SoA, it does not provide a clear mechanism for integrating the various sources of information involved in the agency.

The third theory, the Bayesian integration theory (Synofzik et al., 2009), was developed to address the limitations of the comparator model and the retrospective theory. This theory aims to provide a comprehensive framework that integrates various sources of information from both sensorimotor and cognitive levels to better explain the emergence of the SoA. The Bayesian integration theory builds on the concept of Bayesian inference, a probabilistic method for updating beliefs based on new evidence. According to this theory, the brain combines prior knowledge (prior beliefs) with new evidence (sensory input) to generate updated beliefs about the likelihood of experiencing a SoA. The computational framework of the Bayesian integration theory consists of two levels: the sensorimotor level and the cognitive level. The sensorimotor level includes elements related to actions and action outcomes, such as motor commands, efference copies, and sensory feedback. The cognitive level includes elements related to intentions, expectations, and inferences, such as beliefs about one's ability to cause specific outcomes or the likelihood of a particular outcome occurring in the environment. In this framework, the likelihood of experiencing a SoA is computed using Bayes' rule at each level, taking into account the variability of the result distributions. The formula for this theory is:

$$p(\text{self}|\text{outcome}) = p(\text{outcome}|\text{self}) \times p(\text{self}) / p(\text{outcome})$$

where:

- p(self|outcome) represents the probability that the self perceives a SoA given the outcome.
- p(outcome|self) refers to the probability that the self could cause the outcome.
- p(self) indicates the overall probability of the self as an agent in the environment.

- p(outcome) identifies the probability of the outcome occurring in the environment.

The Bayesian integration theory stands out for its holistic approach, which allows for the combination of multiple information sources, spanning both sensorimotor and cognitive levels. It surpasses the comparator model and retrospective theory by integrating different information sources and adjusting beliefs about the SoA based on new evidence. The comparator model falls short by primarily focusing on prediction errors or sensorimotor incongruence, disregarding temporal aspects of the SoA, and failing to provide an integrative mechanism. Although the retrospective theory addresses the temporal aspect with cognitive inferences based on premotor signals and delayed reafferent sensory information, it still lacks a clear framework for combining various sources of information involved in the agency. The Bayesian framework, in contrast, computes the likelihood of experiencing SoA using Bayes' rule at both the sensorimotor and cognitive levels, accommodating the variability of result distributions. This framework's strength lies in its ability to account for complexities in the SoA's emergence, such as variations in prior beliefs, uncertainty in sensory input, and the influence of contextual factors.

How to Measure Sense of Agency

It has always been difficult to directly measure and study subjective phenomena such as SoA, and many others. However, when it comes to SOA, there are three primary approaches for measuring it.

The first approach is Intentional Binding. IB refers to the perceived temporal compression between a voluntary action and its outcome, in which the time between a voluntary action and its consequence is viewed as shorter than the time between two physically similar involuntary actions (Haggard et al., 2002). This paradigm is typically measured using a clock, which makes it be considered the implicit way to measure SoA. Participants are not directly asked whether they experienced agency or if they believe they caused the action and its consequences. Instead, they are asked to report the timing of a voluntary action (typically a keypress) onset and/or the subsequent event (typically a tone) onset (Haggard et al., 2002; Moore & Obhi, 2012). It has been found that participants' judgments of time differ depending on whether their action (keypress) caused the tone (Moore & Obhi, 2012). The results in such studies are usually interpreted as the following: The greater the binding is being reported in the participants' answers, the stronger the SoA (Haggard et al., 2002; Moore & Obhi, 2012).

The second approach to be discussed is Explicit Agency Judgments. In this approach, participants are asked to make explicit judgments about their SoA following an action, this is why it is considered to be the explicit measure of agency (Moore, 2016). For example, participants are asked directly to report Yes or No about their agentic experience e.g., “Yes it was me, or No it was not me”. Typically, the participants perform an action but do not see it directly. They are instead provided with some type of feedback on a screen. This input may reflect the participant's action or the action of someone or something else (maybe an experimenter or a computer), and the participants are asked to identify whose movement it is or if they caused the action, rated on a scale. Importantly, the researcher ensures that the action being displayed for the participants is uncertain (covering the hand that they are using in the experiment with a towel) and sometimes a delay could be inserted between the action and the feedback presented to the participant (Daprati et al., 1997; Farrer et al., 2008; Moore, 2016; van den Bos & Jeannerod, 2002). This makes the participants reflect on what is happening and then infer and determine if they caused the action and its consequences.

The third approach to be discussed is Sensory Attenuation. Sensory attenuation is a phenomenon in which the experience of sensory effects of one's own activities is decreased when contrasted to sensory stimuli provided by others. The brain's ability to forecast the sensory consequences of self-generated activities is assumed to be the source of this attenuation, allowing it to filter out or decrease the sensory effect of those outcomes (Shergill et al., 2003). As a result, self-generated sensations are evaluated as less vivid or conspicuous than feelings generated by other sources. In a classic experiment, participants are asked to tickle themselves and rate the ticklishness of the sensation, which is then compared to the rating when another person tickles them. This difference in perceived ticklishness is attributed to sensory attenuation, as the brain can predict the sensory consequences of self-tickling but not those of being tickled by another person (Blakemore et al., 1998). Sensory attenuation is an effective experimental paradigm for examining the SoA due to how it directly tests the brain's capacity to discriminate between self-generated and externally produced experiences. Researchers may gain insights into the elements that impact the SoA and the underlying cognitive and neurological mechanisms by evaluating the degree of sensory attenuation in diverse settings.

Neural Basis of Sense of Agency

Recent neuroimaging studies have helped to locate the brain regions that carry out the numerous cognitive and computational processes that underpin the SoA. The comparator model, in particular, has prompted research on brain areas connected with sensory incongruence. For example, researchers compared brain activity between congruent action (strong agency) and incongruent action (weak agency) (David, 2012; Haggard, 2017). E.g., if the direction of a cursor movement matches the direction in which participants move a joystick, this would be the congruent condition, and if the cursor moved in the opposite direction of the joystick, this would be the incongruent condition (Farrer & Frith, 2002). Results have consistently highlighted the role of the parietal cortex in SoA, including the involvement of the inferior parietal and temporoparietal junction in incongruent conditions compared to congruent conditions. The role of the supplementary motor area (SMA), located in the frontal cortex, has also been emphasized (Haggard, 2017; Seghezzi et al., 2019; Spengler et al., 2009; Sperduti et al., 2011; Zito et al., 2020). Furthermore, the bilateral precentral gyri and the left inferior parietal lobe were sensitive to prediction error (sensorimotor incongruency) (Wen & Imamizu, 2022).

Interestingly, Moore and colleagues (2010) used continuous theta burst stimulation to investigate the neural basis of IB on two target sites: the pre-SMA and the primary motor cortex (M1). The effect of stimulation of these target areas was compared to that of stimulating a control region, the sensory leg area. According to the findings of this study, pre-SMA is likely to play a significant part in IB, and its job may be to pre-emptively link intents to the sensory consequences of action (Moore et al., 2010).

Other studies suggest that in the absence of voluntary movement, the temporoparietal junction responds to unexpected external sensory stimuli (Kincade et al., 2005). As a result, its activation in non-agency situations may not represent the process of attributing agency but rather one possible outcome of that process (namely, the judgment that an event is externally caused). More specifically, the right supramarginal gyrus (SMG), the angular gyrus, and medial and lateral prefrontal areas such as the left inferior frontal gyrus (IFG) were particularly sensitive to subjective (explicit) judgments of agency and were associated with non-agency over outcomes (Farrer et al., 2003; Haggard & Chambon, 2012; Sperduti et al., 2011). In an attempt to link the theoretical models of the SoA to these findings, it could be noticed that the comparator model has gained support from studies that showed brain activity in the parietal cortex. However, the retrospective approach has received support from studies that have reported brain activity in SMG and IFG.

Joint Actions

Humans are social creatures, and they rarely operate alone, rather, they are continuously interacting with and coordinating their activities with others around them. As a result, joint action is defined as any kind of social interaction in which two or more humans coordinate their activities in space and time to effect a change in the environment (Sebanz et al., 2006). In joint action tasks, little is known about how the presence of co-actors affects the SoA. According to some recent studies, when two human actors participate in a task in which one of them takes an action that results in an effect, only the person who initiates the effect experiences an explicit SoA. However, both human actors show significant IB (Obhi & Hall, 2011; Strother et al., 2010). Hence, when two individuals participate in a joint action environment, a new agentic identity (a 'we' identity) is formed automatically at the pre-reflective level, even though their subjective experience of agency differs, as does their participation in creating the outcome. Yet, little is known about the effects of replacing a human actor with an automated tool or robot on our SoA. As a result, the new aspect of joint action (Human-Robot Joint action & human-AI interaction) will be the focus of this systematic review.

Automation

Automation is generally defined as the use of technology to perform tasks without human intervention or input. Although humans may be present as spectators, observers, or even participants, the automated process functions autonomously (Groover, 2020). While simple automation can be purely mechanical or based on pre-set instructions, advanced automation often leverages what is known as artificial intelligence (AI). This incorporation of AI into automation reflects our deep interest in replicating human-like thought and decision-making. We have always sought to understand the intricacies of our cognition—how our brain, a compact mass of matter, can perceive, comprehend, predict, and influence an environment that might be more complex than the brain itself.

As we delve deeper into the intricacies of human cognition, the field of AI propels us further. It aims not just to understand the genesis of thought and decision-making but also to create intelligent entities capable of independent decisions (Russell & Norvig, 2020). The term AI was presented for the first time in 1956 in the U.S. and the first AI program was also developed in the same year by Allen Newell & Herbert Simon. Its purpose was to prove logical theorems and it was based on symbol manipulation (Newell & Simon, 1956). Since then, numerous programs, devices, and tools have started to become automated with different levels of automation embedded in each of them. For example, according to the Society of Automobile Engineers International (*SAE J3016 Automated-Driving Graphic*, 2019), there are six stages of automation in vehicles such as cars and buses, among others. It starts with SAE level 0: There is no automation; the human driver is fully responsible for the performance of the driving task. SAE Level 1: Driver assistance, a driver assistance system that allows the automobile to steer or accelerate/decelerate by utilizing information about the driving environment, while the human driver does all other parts of the driving task. SAE Level 2: Partial Driving Automation, several driver assistance systems enable the automobile to execute steering and acceleration/deceleration utilizing data from the driving environment, while the human driver oversees the other components of the driving task. SAE Level 3: Conditional Driving Automation, the driving mode-specific performance of all aspects of the driving task by an automated driving system, while the human driver is fully responsible for responding appropriately to any request for intervention. SAE Level 4: High Driving Automation, the performance of all aspects of the driving task by an automated driving system under limited conditions, with the human driver not required to respond or take over at any point. SAE Level 5: the same as Level 4, but the automobile may now drive anywhere in any condition (*SAE J3016 Automated-Driving Graphic*, 2019).

Neural Basis of Tool Use

Humans have used tools for thousands of years, such as a wooden stick, a knife, a hammer, and so on. Each of these tools might serve one or many purposes. However, during the last several decades, tools have begun to develop substantially, resulting in automated tools which could be termed the next generation of tools. Automated tools can function without human involvement or input, and the effects of interacting with such tools have not been thoroughly investigated. However, using "ordinary tools" with no automation embedded in them has been well documented, which is why results and findings from research on "ordinary tools" will be used as a reference here to understand how the brain receives, processes, and acts when people use such tools in general.

When discussing tool use, two levels of processing should be addressed: the conceptual level (tool knowledge) and the production level (real tool use) (Lesourd et al., 2021). Research of the neurocognitive basis of tool usage has been the focus of intensive studies over the last decades and has been largely inspired by observations of individuals with left brain damage. A lesion in the left hemisphere could cause apraxia, a cognitive impairment of motor control, which affects object-related activities by deteriorating the representations that enable these movements (Goldenberg & Hagmann, 1998; Lesourd et al., 2021).

The conceptual level can be broken down into three sublevels, 1) Sensorimotor knowledge 2) Mechanical knowledge 3) Semantic knowledge (Baumard et al., 2014).

The sensorimotor knowledge sublevel includes information on tool manipulations and the movements involved with the normal manipulation of a certain tool (Baumard et al., 2014). Typically, injury to the left inferior parietal lobe may hinder this type of information (Baumard et al., 2014; Buxbaum & Saffran, 2002). Furthermore, different brain areas are activated when people think about manufactured objects such as tools in general rather than just natural ones (e.g., cars vs. rocks). Chao et al., (2002) reported that when participants were shown pictures of tools rather than natural objects, the left ventral premotor cortex was active. This research shows that created objects, such as tools, are processed differently by the brain than natural ones, such as rocks.

The mechanical knowledge sublevel encompasses an understanding of the connections between the physical attributes of tools and the objects they interact with (e.g., hammering requires that the hammer is heavier than the nail) (Baumard et al., 2014). Damage to the left inferior parietal lobe, a brain region crucial for processing sensory information, spatial awareness, and tool use, can impede the application of this mechanical knowledge (Goldenberg & Spatt, 2009). As a result, individuals with damage to this area may struggle with tasks requiring an understanding of mechanical properties and tool-object interactions, leading to difficulties in effectively using tools or performing tasks involving tool manipulation.

The semantic knowledge sublevel, as described by Baumard et al. (2014), includes information about the typical relationships between recognized tools and their associated objects, as well as the context in which they are commonly used (e.g., understanding the connection between a hammer and a nail, and recognizing that they are typically used in construction or woodworking). Damage to the left anterior temporal lobe can impair an individual's ability to access and utilize this semantic knowledge, leading to difficulties in identifying appropriate tools for specific tasks or understanding how a tool should be employed in a particular context.

The production level is responsible for developing movement patterns while taking both environmental constraints and the tool action representation built by the conceptual system into consideration (Baumard et al., 2014; Osiurak, 2013). The dorsal stream regions are believed to be associated with motor representations related to familiar tools and their use (Baumard et al., 2014; Culham & Valyear, 2006). Specific visuomotor activities, such as reaching, grasping, and eye movements during tool use, are governed by specialized areas within the parietal lobe. Furthermore, the human parietal cortex is engaged in processing and perceiving action-related information even in the absence of explicit action performance (Culham & Valyear, 2006), highlighting its importance in coordinating tool use and motor planning.

Aim of the Present Thesis

The present thesis aims to discern how automated tools influence human SoA, and if tools with different levels of automation could affect SoA differently. Building upon the burgeoning field of "Human-Robot Joint Action" or "human-AI interaction," this work explores the relationship between automation levels in tools and the impact on human SoA. In contrast to previous research that has remained largely theoretical, this study extends the inquiry by conducting a systematic review of empirical evidence. Moreover, while most prior studies have focused on a single tool or task, this research adopts a broader approach, examining how automation affects SoA across a range of tools and tasks.

Unveiling the relationship between automation and human SoA may enhance our understanding of how we interact and co-exist with increasingly autonomous tools. This is crucial in an era where technology continues to permeate every facet of our lives. Furthermore, it may reveal novel insights into the psychological impacts of automation, thus informing the design and implementation of future automated systems.

Methods

Literature Search

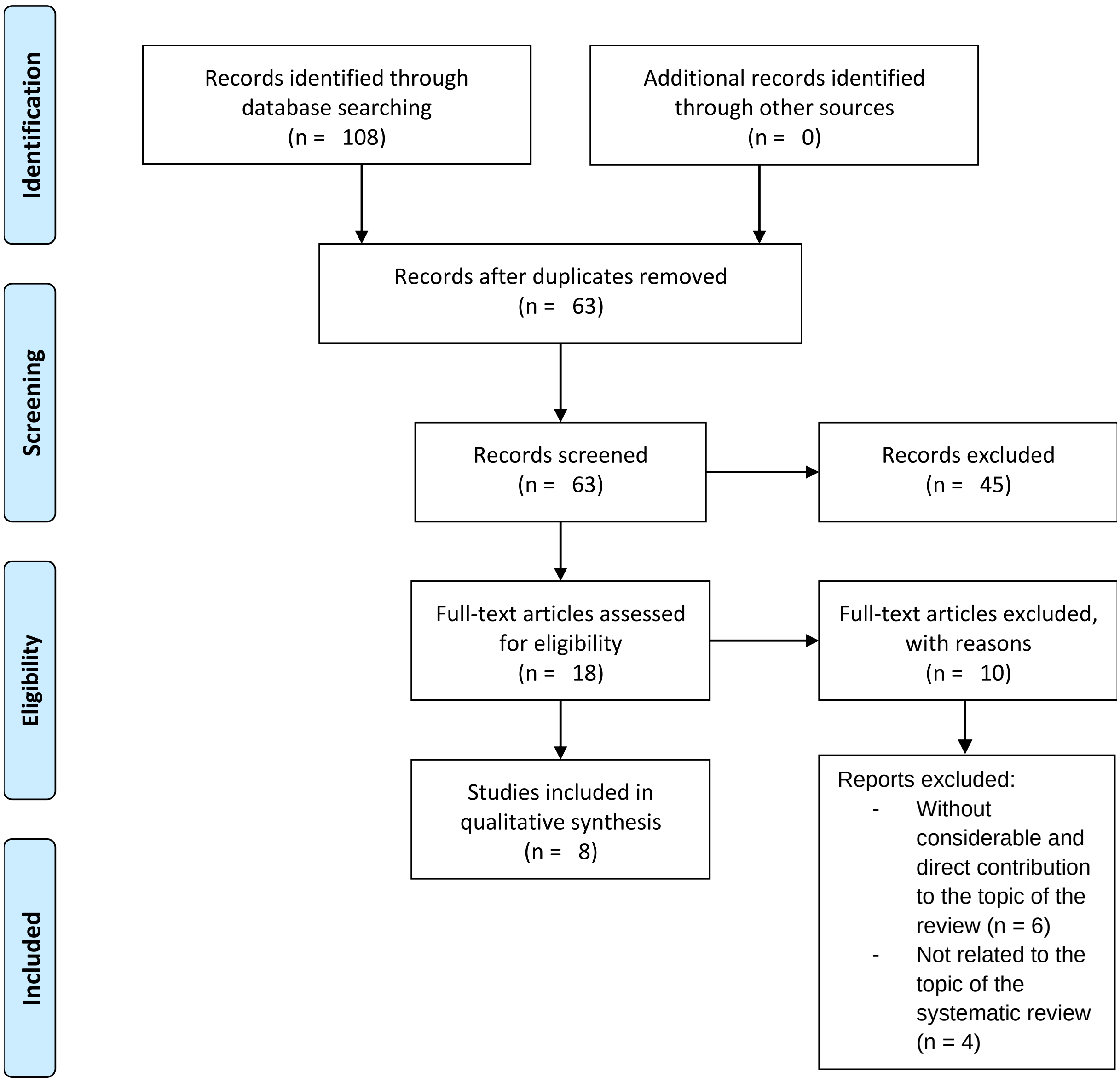
A literature search in two databases, Scopus and MEDLINE EBSCO, was conducted on July 5, 2022. The timespan was not set to any specific date, and each database was searched using the following keyword strings: ("Sense of Agency" AND Automation*), ("Sense of Agency" AND "Human-Robot Joint action"), ("Sense of Agency" AND "Human–AI interaction") and (autom* AND "Sense of Agency"), without applying any filter to the search process. All results were imported into the online software Rayyan (Ouzzani et al., 2016). A total of 108 articles were identified. Using Rayyan's duplicate removal function, duplicates were removed (n=45). During the first screening, the titles and abstracts of each article were examined, resulting in the exclusion of an additional 45 articles. The second screening was based on reading the full text of each article. This led to the exclusion of 10 articles due to irrelevant study design and incorrect study type. Ultimately, eight articles remained for further analysis (See Figure 1).

Inclusion and Exclusion Criteria

Articles were included if they were experimental / empirical studies published in a peer-reviewed journal and written in the English language. Due to the limited number of articles found and the primary focus of this systematic review (i.e., discover the effects of automated tools on the SoA), all studies that used any type of automated tool were included without regard to a specific tool nor to the level(s) of automation. Moreover, the studies that have investigated only SoA, or Human-Robot Joint action, or Human–AI interaction, or Automation/Automated tools in a separate manner without looking at the relation between them were excluded. Furthermore, articles that investigated what interface to be used in different levels of automation were excluded.

Figure 1

PRISMA 2009 Flow Diagram (Moher et al., 2009): Standard Flow Diagram used to Document the Literature Search Process.


Identification
Screening
Eligibility
Included
Records identified through database searching (n = 108)
Additional records identified through other sources (n = 0)
Records after duplicates removed (n = 63)
Records screened (n = 63)
Records excluded (n = 45)
Full-text articles assessed for eligibility (n = 18)
Full-text articles excluded, with reasons (n = 10)
Studies included in qualitative synthesis (n = 8)
Reports excluded:
- Without considerable and direct contribution to the topic of the review (n = 6)
- Not related to the topic of the systematic review (n = 4)

Results

A total of 8 studies were included in the systematic review, involving 278 participants. The characteristics of the included studies are summarized in Table 1 and Table 2.

Four studies (Nakashima & Kumada, 2020; Tamura et al., 2013; Wen et al., 2021; Yun et al., 2017) employed two levels of automation. The first level represented a full control condition, where participants maintained complete control over task initiation, planning, and execution. The second level constituted an automated condition, in which the utilized tools or programs exerted significant influence over the task.

In the study by Tamura et al. (2013), 21 participants were involved in a series of three experiments that manipulated their activeness, SoA, and predictability of the experiment's conditions. The first experiment engaged participants in both active and passive conditions, where the participants controlled or observed the movement of a searchlight, respectively. The reaction times measured showed a significant difference between the active ($M = 435.50$ ms, $SD = 35.73$) and passive ($M = 464.99$ ms, $SD = 41.43$) conditions, revealing that activeness can improve human cognitive performance ($p < 0.01$). The second experiment examined the relationship between a SoA and cognitive performance by introducing five different temporal delay conditions between joystick input and searchlight motion. The degree of the SoA, represented as the "YES ratio", significantly decreased as the delay increased ($p < 0.01$). Moreover, there was a negative correlation ($r = -0.46$) between the degree of the SoA and reaction time, suggesting that a higher SoA leads to better cognitive performance. In the third experiment, the predictability of the searchlight's movement was manipulated. The searchlight was moved automatically and participants were subjected to high and low predictability conditions, which significantly influenced reaction times. The average reaction time under the high predictability condition was significantly shorter ($M = 454.91$ ms, $SD = 29.42$) than that under the low predictability condition ($M = 477.64$ ms, $SD = 26.20$) ($p < 0.05$).

In the studies by Yun et al. (2017) and Wen et al. (2021), participants engaged in driving tasks using a laboratory simulator. Yun et al. (2017) conducted a study in which participants were subjected to simulated driving tasks under various conditions. Participants' SoA was assessed after each trial using a 7-point scale. The SoA ratings, reported as mean±SD, for the licensed participants under the self-control condition was 5.07 ± 0.96, under the automated condition was 3.79 ± 1.90, and under the delayed condition was 3.31 ± 1.20. The SoA rating was significantly lower in the delayed condition compared to the self-control condition ($p < 0.01$). However, the decrease in the automated condition was not significant ($p = 0.13$). A participant who did not have a driver's license provided an intriguing case in which the SoA was greater under the automated condition than under the self-control condition. The mean SoA results were as follows: under the self-control condition was 4.33 1.25, under the automated condition was 4.33 2.36, and under the delayed condition 3.33 0.47. Electroencephalography data were gathered to determine the relative strength of the alpha band, however, due to technological challenges, only data from one licensed participant were evaluated. The mean relative power of the alpha band varied across conditions. Electrodes Cz, C2, CP6, P3, Pz, and P4 demonstrated higher mean relative power in the automated or delayed condition as compared to the self-control condition, while C1 demonstrated an opposite pattern. However, there was no statistically significant difference in power between self-control and delayed situations (ANOVA, $p > 0.05$).

In a study by Wen et al. (2021), 19 participants were tasked with controlling a vehicle in both manual and assisted driving modes under high and low-risk conditions. Self-reported agency ratings did not significantly vary based on the driving mode or risk level (all ps > .05). However, subjective performance ratings did significantly vary based on both the driving mode ($p = .026$) and the risk level ($p < .001$). In the analysis of actual performance based on the distance coefficient of variation (DCV) and speed coefficient of variation (SCV), the DCV was significantly influenced by the risk level ($p = .018$) but was not significantly affected by the driving mode ($p = .122$). In contrast, the SCV was significantly affected by both the driving mode ($p < .001$) and the risk level ($p < .001$). In summary, the SoA remained consistent across all conditions. However, subjective performance ratings and actual performance measures (DCV and SCV) significantly varied depending on the driving mode and risk level.

In the study by Nakashima and Kumada (2020), 27 participants were involved in an experiment where they had either full control or no control over halting a moving object. Participants rated their SoA on a 100-point scale following each trial. The study found significantly smaller decreases in SoA in the gradual-stop condition compared to the sudden-stop condition ($p < .001$). Also, the decrease in SoA was smaller in the goal-present condition compared to the goal-absent condition ($p < .001$). However, the interaction of these factors was not statistically significant ($p = .38$). Participants rated their SoA higher in manual control trials compared to automatic control trials ($p < .001$). The interaction between control type and type of stop was significant ($p < .001$), as was the interaction between control type and goal ($p < .001$).

Two out of the eight included studies featured three levels of automation. In the first study conducted by Di Plinio et al. (2019), 60 participants engaged with three levels of automation: positive-control, no-control, and negative-control. Participants were asked to judge the timing between a keypress that initiated a rotating clock hand on a screen and subsequent feedback via a tone. The study found a significant predictive component in the positive-control group ($p = .01$) but not in the no-control group ($p = .77$) or the negative-control group ($p = .29$). A notable difference in the predictive component was found between the positive-control and negative-control groups ($p = .03$), and between the positive-control and no-control groups ($p = .02$). However, no significant difference was observed between the no-control and negative-control groups ($p = .55$). The retrospective component was not found to be significant in any of the experimental conditions.

The second study, by Zanatto et al. (2021), involved 74 participants who were instructed to judge the interval between a keypress and an acoustic tone. The three levels of automation were Human Decision condition, System Warning condition, and System

Decision condition. In the Human Decision condition, participants pressed the spacebar at any time after a fixation cross appeared on the screen. In the System Warning condition, a warning signal reminded participants to press the spacebar after the appearance of the fixation cross. The reminder and the fixation cross remained on the screen until the spacebar was pressed. In the System Decision condition, participants were asked not to press the spacebar and wait for the computer to perform the keypress. The results showed a significant main effect of the automation level on the estimation score ($p < .001$ for all conditions). In the Human Decision condition, the estimation score was the shortest mean = -16.90 ms (SD = 53.53 ms). In the System Warning condition, the estimation score was 57.22 ms (SD = 42.12 ms), and in the System Decision condition, the estimation score was 93.04 ms (SD = 59.04 ms).

Berberian et al. (2012) conducted a study with 12 participants, focusing on four levels of automation named Full Operator Control (FOC), Automatic Decision and Implementation Operator Engagement (AD-OIE), Automatic Decision and Implementation - Operator Engagement (ADI-OE), and Full Automatic Control (FAC). Participants were tasked with monitoring and possibly intervening in an aircraft's progress following a predefined flight path displayed on a screen. The level of participant involvement varied across each level of automation. Results showed a significant main effect of Action/Effect delay ($p < .01$), indicating that participants could effectively track the physical variation of the interval. The main effect of Automation level was also significant ($p < .01$), with post-hoc analysis revealing that interval estimates increased in relation to the level of automation. This implies that as the system became more automated, the perceived action-effect interval lengthened, indicating a decrease in the SoA. Moreover, the two-way interaction between Automation Level and Action/Effect delay was significant ($p < .01$), with interval estimates from the medium and large Action/Effect delays strongly modulated by the automation level ($p < .01$). However, estimates from the small Action/Effect delays were not significantly modulated ($p > .01$). Additionally, the explicit judgment of agency was influenced by the Automation Level ($p < .01$). As the level of automation increased, the judgment of causality decreased monotonically (all p's < .01). Lastly, a significant negative correlation was found between authorship and binding effects (Mean $r = -0.84$, $SD = 0.105$, $p < .001$). This suggests that as actual levels of control varied, changes in intentional binding closely tracked explicit judgments of agency. For further statistical details, see the original study.

Finally, in Ueda et al.'s (2021) study, two experiments were conducted with 60 participants tracking a moving object using a joystick-controlled cursor. In Experiment 1, there were 6 various levels of automation, and they were complete control, 33%, 66%, -33%, -66%, and full automation. The self-reported SoA ratings showed an increase with assistive automation, plateauing after 33% automation, and then declining in the full automation condition. One-way repeated ANOVA tests showed significant effects of task conditions on both control ratings ($p < 0.01$) and performance ratings ($p < 0.01$). In Experiment 2, there were also 6 various levels of automation, and they were complete control, 80%, 85%, 90%, 95%, and full automation. The control ratings indicated a decrease in SoA when automation exceeded 90%. The one-way repeated ANOVA tests demonstrated significant effects of task conditions on both control ratings ($p < 0.01$) and performance ratings ($p < 0.01$). Furthermore, regression analyses showed a positive relationship between subjective evaluations of apparent tracking performance and SoA from -66% to 66% automation. However, a weak negative relationship was observed between subjective evaluations of apparent tracking performance and SoA at automation levels of 80% and above. Overall, the study found that SoA was enhanced by the increase in the level of automation, but it began to decline when the level of automation exceeded 90%. For further statistical details, see the original study.

Table 1

Summary of Study Characteristics for The Included Studies in This Systematic Review.

Author(s)	Year	Country	Study Design	Task	Sample Size	Gender Distribution (M/F/NB)	Mean Age (SD)
Berberian et al.	2012	United States	Within-subject	Supervision task with different autopilot settings of an aircraft's simulator	13	9 M / 4 F / 0 NB	32 (-)
Di Plinio et al.	2019	Italy	Within-subject	A key press task.	60	28 M / 32 F / 0 NB	24 (3)
Nakashima & Kumada	2020	Japan	Within-subject	A key press task.	27	12 M / 15 F / 0 NB	21 (-)
Tamura et al.	2013	United Kingdom	Within-subject	A search using a joystick.	21	12 M / 9 F / 0 NB	22.5 (-)
Ueda et al.	2021	Germany	Within-subject	A tracking task using a joystick.	60	32 M / 28 F / 0 NB	21.8 (3.1)
Wen et al.	2021	China	Within-subject	Drive a car in a laboratory simulator.	19	19 M / 0 F / 0 NB	22.4 (1.4)
Yun et al.	2017	South Korea	Within-subject	Drive a car in a laboratory simulator.	5	5 M / 0 F / 0 NB	- (-)
Zanatto et al.	2021	Brazil	Within-subject	A key press task.	74	32 M / 37 F / 5 NB	21.92 (4.18)

Note. This description indicates that the table contains a summary of the study characteristics for the studies included in the systematic review. For the Gender Distribution, M stands for Male, F stands for Female, and NB stands for non-Binary.

Table 2

Summary of Empirical Studies on the Impact of Automation Levels on Human Performance and Sense of Agency.

Study ID	Author(s)	Variables / Levels of automation	Results
1	Berberian et al. 2012	Four levels: Full Operator Control (FOC), Automatic Decision and Implementation Operator Engagement (AD-OIE), Automatic Decision and Implementation - Operator Engagement (ADI-OE), and Full Automatic Control (FAC)	Automation level significantly affects perceived duration of intervals between actions and effects [$F(3,36) = 26.154$; $p < .01$, $\eta p2 = .69$], and explicit judgement of agency [$F(3,36) = 46,204$; $p < .01$, $\eta p2 = .79$]. Level of automation is inversely related to SoA (Mean $r = -0.84$).
2	Di Plinio et al. 2019	Three levels: Positive-control, No-control, Negative-control	The predictive component is significant in the positive-control group [$F(1, 228) = 6.8$, $p = .01$] but not in the no-control or negative-control groups. Differences are observed between positive-control and the other two groups.
3	Nakashima & Kumada 2020	Two levels: Full control, No control	Decrease in SoA observed in certain conditions, including manual control trials vs automatic control trials [$F(1, 26) = 193.50$, $p < 0.001$, $\eta p2 = 0.88$].
4	Tamura et al. 2013	Two levels: Active control, Passive control	Active control leads to shorter reaction times ($M = 435.50$ ms) than passive control ($M = 464.99$ ms), and decreased degree of agency as delay increased. [bootstrap paired t-test, $p < 0.01$].
5	Ueda et al. 2021	Five level: Complete control, 33%, 66%, -33%, -66%, Full automation Five level: Complete control, 80%, 85%, 90%, 95%, Full automation	Significant effects of task conditions on control ratings [$F(5, 145) = 11.70$, $p < 0.01$, $\eta p2 = 0.29$] and performance ratings [$F(5, 145) = 118.67$, $p < 0.01$, $\eta p2 = 0.80$] observed. Evaluations of tracking performance and SoA vary depending on automation levels, with a shift in the relationship at high automation levels (80% and above).
6	Wen et al. 2021	Tow level: Full control, Assisted	No significant difference in SoA ratings between conditions. However, subjective performance rating varies depending on driving mode [$F(1, 18) = 5.909$, $p = 0.026$, $\eta 2p = 0.247$] and risk level [$F(1, 18) = 30.106$, $p < 0.001$, $\eta 2p = 0.626$].

7	Yun et al. 2017	Two level: Full control, Fully automated	Licensed drivers show lower SoA ratings in automated condition (Mean Automated = 3.79) compared to full control (Mean Full control = 5.07). No significant difference observed among unlicensed drivers.
8	Zanatto et al. 2021	Three level: Human Decision, System Warning, System Decision	Automation level significantly affects estimation score, with System Decision automation leading to highest score (93.04 ms).

Note. SoA refers to Sense of Agency. FOC refers to Full Operator Control, AD-OIE refers to Automatic Decision and Implementation Operator Engagement, ADI-OE refers to Automatic Decision and Implementation - Operator Engagement, and FAC refers to Full Automatic Control. p refers to p-value. $\eta p2$ refers to partial eta-squared, a measure of effect size in ANOVA studies. The Mean r refers to mean correlation coefficient.

Discussion

The current systematic review investigated the effect of various levels of automation on human SoA. Overall, the results suggested that when a tool/device/program becomes fully automated, humans experience a decrease in SoA and, in some situations, a weakened outcome, such as a longer time required to react with the task. A possible explanation here is the neural overlap of SoA and tool-use. Both these processes involve similar brain regions, such as the inferior parietal cortex and the pre-SMA. For example, in SoA, the pre-SMA and the temporo-parietal junction, are associated with control of actions (Haggard, 2017; Seghezzi et al., 2019; Spengler et al., 2009; Sperduti et al., 2011; Zito et al., 2020). Moreover, the Inferior Parietal Lobe is sensitive to sensorimotor incongruency in SoA (Farrer et al. 2008; Wen & Imamizu, 2022).

This is mirrored in tool use, where the parietal cortex is engaged in processing and perceiving action-related information even in the absence of explicit action performance and is key to coordinating tool use and motor planning (Maravita & Iriki, 2004; Culham & Valyear, 2006). Damage to the left inferior parietal lobe can impede mechanical and sensorimotor knowledge, affecting the ability to understand the physical attributes of tools and their interactions with objects and the normal manipulation of tools (Baumard et al., 2014; Buxbaum & Saffran, 2002; Goldenberg, 2009).

Another possibility to explain why humans may experience a decrease in SoA is that they begin to experience "out-of-the-loop" performance when the tool is in control of the work at hand. The "out-of-the-loop" performance issue means that manipulators believe they are not using the tool (or the system) at all, and hence they are not participating in the work or action. Instead, they are now working on a new duty, which is the monitoring of the original task.

To maintain SoA, one could add an objective to the task. Wen et al. (2021) suggested that if automated decisions have the same goal/intention as the manipulators/users, it may be able to eliminate the trade-off between SoA and task performance while also enhancing the task performance. Adding an overarching goal to a task might help manipulators/users to have a feeling of accomplishment when they complete a task by themselves or with the assistance of an automated tool. The findings of Nakashima and Kumada (2020), Tamura et al. (2013), and Wen et al. (2021), which focus on the examination of cognitive processes and intentions, align with the retrospective theory proposed by Wegner et al. (2004) and Wegner and Wheatley (1999). Furthermore, these empirical findings support the idea that SoA is a derivative of cognitive judgments, which are based on the integration of temporally diverse data (Synofzik et al., 2013). These findings support the idea presented by Wegner (2003) and Wegner et al. (2004) that cognitions and mental representations play a role in determining actions.

On the other hand, the findings of Yun et al. (2017) suggested a very interesting point. Participants with a driver's license reported a decrease in SoA when the automated system took over the task. Interestingly, the one driver without a driver's license reported an increased SoA in the automated condition compared to the self-control condition. Thus, inexperienced drivers may feel enhanced SoA under driving assistance. Such findings could be explained with the comparator model (Frith, 2005; Frith et al., 2000) that posits that the SoA arises from the brain's ability to generate sensory predictions based on motor commands and compares these predictions with actual sensory feedback. In the instance of the inexperienced driver, he/she lacked sufficient sensory predictions of the task's outcome (in this case driving). As a result, he/she was unable to link any of the outcomes to any sensory predictions, necessitating the need for some outside assistance to finish the task. The assistance offered by the automated system enabled the participants to achieve their objective and complete the task, allowing them to have a feeling of accomplishment at the end. As a result, the SoA increased.

As we start moving from the two levels of automation to the three or four levels of studies, it becomes apparent that the SoA does not plummet abruptly between the fully manual control condition and the completely automated one. Instead, the SoA appears to adjust progressively in line with the degree of automation. Specifically, there is a decrease in the SoA commensurate with the increasing level of automation. This phenomenon was markedly evident in research by Zanatto et al. (2021), which demonstrated a steady increase in the perceived interval between action and effect. This perceived elongation of the action-effect interval suggests a diminution in intentional binding, which can be understood as a gradual attenuation of the SoA. Those findings align with the findings of the studies by Di Plinio et al. (2019) and Berberian et al. (2012). These studies demonstrated a similar trend: as the level of automation in a tool or system rises, participants experience a corresponding decrease in the SoA. In addition, Berberian et al.'s (2012) study yielded consistent results from both the implicit method of assessing agency (via intentional binding) and the explicit approach of measuring the SoA (through explicit reporting). This consistency between the two methodologies not only strengthens the significance of the findings but also suggests a harmonious relationship between the two different approaches to measure SoA.

Finally, the results of the study by Ueda et al. (2021) come with interesting findings. In this experiment, it is noticed that users/manipulators SoA was enhanced by increasing automation but began to decline when the level of automation exceeded 90%. Above this level, participants started to lose their SoA. This could lead us back to the first discussion point. When participants start to experience that the tool begins to take over and they are taken out of the control loop, they might feel that they are not actively using the tool anymore, and they have a passive supervision duty over the main task. This could lead them to lose their SoA. On the other hand, growing evidence suggests that providing users/manipulators with opportunities for continuous operation and assisting them in improving their work performance through automation may increase their SoA. Automation may be beneficial for users with an adequate level of expertise for the task at hand. However, this topic is still in its infancy phase, and more research is needed to understand its implications on SoA and people's lives.

Limitations, Strengths

The strengths of the studies included in this review are several. First, the studies cover a broad spectrum of automation levels, ranging from semi-automated to fully automated systems, thereby providing a comprehensive overview of the impact of automation on SoA. Second, the studies employed both implicit and explicit measures of SoA, offering a more complete picture of the effects of automation. Third, the studies included various types of tasks and tools, increasing the generalizability of the findings.

One of the main limitations is the lack of a standardized measure of SoA across studies, which makes it difficult to compare results directly. Second, while automation levels were varied, the studies did not thoroughly investigate the impact of specific degrees of automation on SoA. Moreover, most studies did not consider the role of individual differences, such as experience and familiarity with the task or tool, which could significantly influence SoA.

Ethical and Societal Considerations

As automation becomes increasingly pervasive in our lives, understanding its impact on the user´s SoA is crucial. If automation decreases SoA, it may lead to a decrease in motivation, job satisfaction, and performance, particularly in professions heavily reliant on automated tools. Additionally, there may be ethical issues to consider when designing and implementing automated systems. For instance, if an automated system fails, who is responsible? Is it the user, who might not have a strong SoA over the system, or is it the designer or operator of the system?

Furthermore, there are societal implications related to equity and access. Automated systems are often expensive and require a certain degree of technological literacy to use effectively. This could potentially widen the gap between those who have access to these technologies and those who do not.

Conclusion

In conclusion, this review sheds light on the effect of automation and its varying degrees on human SoA. As tools become more automated, users experience a diminished sense of control over the task at hand. This necessitates a shift in responsibility, with individuals focusing on supervising and monitoring the tasks performed by the automated tools. However, further research is needed to address the limitations, ethical concerns, and appropriate levels of automation for different tasks. This will be vital for guiding the development and implementation of automated systems in a manner that benefits and ensures fairness for all members of society.

Word count: 7591

References

Baumard, J., Osiurak, F., Lesourd, M., & Gall, D. L. (2014). Tool use disorders after left brain damage. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00473

Berberian, B., Sarrazin, J., Blaye, P. L., & Haggard, P. (2012). Automation technology and sense of control: A window on human agency. *PLOS ONE*, *7*(3), e34075. https://doi.org/10.1371/journal.pone.0034075

Blakemore, S., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640. https://doi.org/10.1038/2870

Blakemore, S., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242. https://doi.org/10.1016/s1364-6613(02)01907-1

Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: dissociations in apraxic and nonapraxic subjects. *Brain and Language*, *82*(2), 179–199. https://doi.org/10.1016/s0093-934x(02)00014-7

Chao, L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. *Cerebral Cortex*, *12*(5), 545–551. https://doi.org/10.1093/cercor/12.5.545

Culham, J. C., & Valyear, K. F. (2006). Human parietal cortex in action. *Current Opinion in Neurobiology*, *16*(2), 205–212. https://doi.org/10.1016/j.conb.2006.03.005

Daprati, E., Franck, N., Guillaume, N., Proust, J., Pacherie, E., Dalery, J., & Jeannerod, M. (1997). Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, *65*(1), 71–86. https://doi.org/10.1016/s0010-0277(97)00039-5

David, N. (2012). New frontiers in the neuroscience of the sense of agency. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00161

Di Plinio, S., Arnò, S., Perrucci, M. G., & Ebisch, S. J. H. (2019). Environmental control and psychosis-relevant traits modulate the prospective sense of agency in non-clinical individuals. *Consciousness and Cognition*, *73*, 102776. https://doi.org/10.1016/j.concog.2019.102776

Farrer, C., Franck, N., Guillaume, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *NeuroImage*, *18*(2), 324–333. https://doi.org/10.1016/s1053-8119(02)00041-1

Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D. J., Inati, S., & Grafton, S. T. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, *18*(2), 254–261. https://doi.org/10.1093/cercor/bhm050

Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, *15*(3), 596–603. https://doi.org/10.1006/nimg.2001.1009

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

Frith, C. D., Blakemore, S., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society B*, *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

Goldenberg, G. (2009). Apraxia and the parietal lobes. *Neuropsychologia*, *47*(6), 1449–1459. https://doi.org/10.1016/j.neuropsychologia.2008.07.014

Goldenberg, G., & Hagmann, S. (1998). Tool use and mechanical problem solving in apraxia. *Neuropsychologia*, *36*(7), 581–589. https://doi.org/10.1016/s0028-3932(97)00165-6

Goldenberg, G., & Spatt, J. (2009). The neural basis of tool use. *Brain*, *132*(6), 1645–1655. https://doi.org/10.1093/brain/awp080

Goswami, A., & Vadakkepat, P. (2019). *Humanoid robotics: A reference*. Springer.

Groover, M. P. (2020). *Fundamentals of modern manufacturing: Materials, processes, and systems*. John Wiley & Sons, Inc.

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

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*, *22*(10), R390–R392. https://doi.org/10.1016/j.cub.2012.02.040

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

Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *The Journal of Neuroscience*, *25*(18), 4593–4604. https://doi.org/10.1523/jneurosci.0236-05.2005

Lesourd, M., Servant, M., Baumard, J., Reynaud, E., Ecochard, C., Medjaoui, F. T., Bartolo, A., & Osiurak, F. (2021). Semantic and action tool knowledge in the brain: Identifying common and distinct networks. *Neuropsychologia*, *159*, 107918. https://doi.org/10.1016/j.neuropsychologia.2021.107918

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

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

Moore, J. W., Ruge, D., Wenke, D., Rothwell, J. C., & Haggard, P. (2010). Disrupting the experience of control in the human brain: pre-supplementary motor area contributes to the sense of agency. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1693), 2503–2509. https://doi.org/10.1098/rspb.2010.0404

Nakashima, R., & Kumada, T. (2020). Explicit sense of agency in an automatic control situation: effects of Goal-Directed action and the gradual emergence of outcome. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.02062

Newell, A., & Simon, H. A. (1956). The logic theory machine--A complex information processing system. *IEEE Transactions on Information Theory*, *2*(3), 61–79. https://doi.org/10.1109/tit.1956.1056797

Obhi, S. S., & Hall, P. (2011). Sense of agency and intentional binding in joint action. *Experimental Brain Research*, *211*(3–4), 655–662. https://doi.org/10.1007/s00221-011-2675-2

Osiurak, F. (2013). Apraxia of tool use is not a matter of affordances. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00890

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). *Rayyan—A web and mobile app for systematic reviews*. *Systematic Reviews*. https://www.rayyan.ai/

Passingham, R. E., & Wise, S. P. (2014). *The neurobiology of the prefrontal cortex: Anatomy, Evolution, and the Origin of Insight*. Oxford Psychology.

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratory behavior. *Journal of Experimental Child Psychology*, *8*(1), 33–39. https://doi.org/10.1016/0022-0965(69)90025-3

Russell, P. R., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

SAE J3016 automated-driving graphic. (2019). SEA. https://www.sae.org/site/news/2019/01/sae-updates-j3016-automated-driving-graphic

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The brain in (willed) action: A meta-analytical comparison of imaging studies on motor intentionality and sense of agency. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00804

Shergill, S. S., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science*, *301*(5630), 187. https://doi.org/10.1126/science.1085327

Spengler, S., Von Cramon, D. Y., & Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *NeuroImage*, *46*(1), 290–298. https://doi.org/10.1016/j.neuroimage.2009.01.047

Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self- and external-agency attribution: a brief review and meta-analysis. *Brain Structure & Function*, *216*(2), 151–157. https://doi.org/10.1007/s00429-010-0298-1

Strother, L., House, K. A., & Obhi, S. S. (2010). Subjective agency and awareness of shared actions. *Consciousness and Cognition*, *19*(1), 12–20. https://doi.org/10.1016/j.concog.2009.12.007

Synofzik, M., Vosgerau, G., & Lindner, A. (2009). Me or not me – An optimal integration of agency cues? *Consciousness and Cognition*, *18*(4), 1065–1068. https://doi.org/10.1016/j.concog.2009.07.007

Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: an interplay between prediction and postdiction. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00127

Tamura, Y., Egawa, M., Yano, S., Maeda, T., Kato, M., & Asama, H. (2013). Activeness improves cognitive performance in Human-Machine interaction. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *17*(3), 425–432. https://doi.org/10.20965/jaciii.2013.p0425

Ueda, S., Nakashima, R., & Kumada, T. (2021). Influence of levels of automation on the sense of agency during continuous action. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-82036-3

Van Den Bos, E., & Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition*, *85*(2), 177–187. https://doi.org/10.1016/s0010-0277(02)00100-2

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences*, *7*(2), 65–69. https://doi.org/10.1016/s1364-6613(03)00002-0

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

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*(7), 480–492. https://doi.org/10.1037/0003-066x.54.7.480

Wen, W., & Haggard, P. (2018). Control changes the way we look at the world. *Journal of Cognitive Neuroscience*, *30*(4), 603–619. https://doi.org/10.1162/jocn_a_01226

Wen, W., & Imamizu, H. (2022). The sense of agency in perception, behaviour and human–machine interactions. *Nature Reviews Psychology*, *1*(4), 211–222. https://doi.org/10.1038/s44159-022-00030-6

Wen, W., Kuroki, Y., & Asama, H. (2019). The sense of agency in driving automation. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02691

Wen, W., Shibata, H., Ohata, R., Yamashita, A., Asama, H., & Imamizu, H. (2020). The active sensing of control difference. *iScience*, *23*(5), 101112. https://doi.org/10.1016/j.isci.2020.101112

Wen, W., Shimazaki, N., Ohata, R., Yamashita, A., Asama, H., & Imamizu, H. (2020). Categorical perception of control. *ENeuro*, *7*(5), ENEURO.0258-20.2020. https://doi.org/10.1523/eneuro.0258-20.2020

Wen, W., Yun, S., Yamashita, A., Northcutt, B. D., & Asama, H. (2021). Deceleration assistance mitigated the trade-off between sense of agency and driving performance. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.643516

Yun, S., Wen, W., An, Q., Hamasaki, S., Yamakawa, H., Tamura, Y., Yamashita, A., & Asama, H. (2017). Investigating the relationship between driver's sense of agency and EEG: Mu-rhythm is more suppressed in higher SoA case. *IEEE Xplore*. https://doi.org/10.1109/mhs.2017.8305264

Zanatto, D., Chattington, M., & Noyes, J. (2021). Human-machine sense of agency. *International Journal of Human-computer Studies*, *156*, 102716. https://doi.org/10.1016/j.ijhcs.2021.102716

Zito, G., Wiest, R., & Aybek, S. (2020). Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLOS ONE*, *15*(6), e0234321. https://doi.org/10.1371/journal.pone.0234321