Received: 17 January 2022 | Accepted: 3 October 2022

DOI: 10.1002/ejsp.2923

RESEARCH ARTICLE

EASP WILEY

A functional analysis of personal autonomy: How restricting 'what', 'when' and 'how' affects experienced agency and goal motivation

Chao Zhang[1] | Supraja Sankaran[2] | Henk Aarts[1]

[1]Department of Psychology, Utrecht University, Utrecht, The Netherlands

[2]Department of Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands

Correspondence
Chao Zhang, Human-Technology Interaction Group, Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, PO Box 513, 5600 MB, Eindhoven, The Netherlands.
Email: c.zhang.5@tue.nl

Chao Zhang is currently affliated with the Human-Technology Interaction Group, Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, but the majority of the work was done when he was working at Utrecht University. Supraja Sankaran is currently affliated with the Department of Communication and Cognition, Tilburg School of Humanities and Social Sciences, Tilburg University, but the majority of the work was done when she was working at Eindhiven University of Technology.

Abstract

Personal autonomy is central to people's experiences of agency and abilities to actively take part in society. To address the challenges of supporting autonomy, we propose a functional model of autonomy, according to which the experience of agency is a function of the opportunity to determine *what* to do, *when* to act and *how* to act in goal-pursuit. We tested the model in three experiments where the three goal-pursuit components could be constrained by another person or an artificial intelligence (AI) agent. Results showed that removing any of the three components from one's own decisions reduced experienced agency (Study 1a and 1b) and lowered motivation to pursue goals in organisational contexts (Study 2). In comparison to the strong and robust main effects, interactions between the components and the effects of the source of restriction (human vs. AI) were negligible. Implications for personal autonomy, algorithmic decision-making and behaviour change interventions are discussed.

KEYWORDS
agency experiences, behaviour change, functional analysis, goal motivation, goal-directed behaviour, personal autonomy

1 | INTRODUCTION

Humans are autonomous agents. They can make unconstrained decisions and act on them to achieve their goals and satisfy their needs. Personal autonomy has distinct effects on human behaviour. Autonomy makes people feel in control of their own behaviour (Borhani et al., 2017; Caspar et al., 2016), facilitates intrinsic motivation to engage in behaviour (Deci & Ryan, 2000), and promotes more sustainable behaviour change (De Young, 1993; Deci & Ryan, 1987). As a more socially reflective construct, autonomy forms the basis for the judgment of responsibility and the understanding of moral behaviour (Bandura, 2002; Tauber et al., 2005). In short, autonomy is central to people's experiences of agency or control, and endows them with the ability to actively take part in large organisations and society as a whole ('citizen empowerment').

While modern societies cherish personal autonomy more and more, supporting and maintaining it become increasingly a challenge. Current societal issues, such as sustainability and public health, require people to change their behaviour. Apart from attempts to alter behaviour through education and incentivisation, behavioural change can be enforced by measures and regulations (Michie & West, 2013), thus creating a tension between respecting individuals' freedom of choice and solving collective problems (e.g., restrictive measures during the COVID-19 pandemic). Similarly, in large organisations, employers' autonomy over their work may be compromised by the employer's goal of maximising the value created by the employees' labour (Kellogg et al., 2020).

Another challenge when considering autonomy restrictions relates to the source of these restrictions. While traditionally autonomy restrictions come from human agents and their institutions, the advances in artificial intelligence (AI) have led to a new source of threat to personal autonomy. In a survey with experts in AI-related fields, autonomy was recognised as a key challenge for the development of AI (Anderson et al., 2018). As smart as AI algorithms may become, the integration of AI-infused systems in human societies depends on whether people accept and trust these systems (Glikson & Woolley, 2020), which closely pertains to issues of autonomy, agency and user control (Väänänen et al., 2021). The threat of AI can be even more salient when AI is used for behaviour change, where machine learning and computational modelling are used to influence behaviour in highly personalised ways (e.g., Spruijt-Metz et al., 2015; van Wissen, 2014; Zhang et al., 2019). It is not hard to imagine a future, where autonomous artificial agents intrude on human decision-making, from what career to pursue to what food to eat for dinner (Sankaran et al., 2020). AI's impacts on human behaviour have already been evident in organisational contexts, particularly in the form of 'algorithmic management' (Kellogg et al., 2020; Schildt et al., 2017), where workers in the so-called 'gig' economy (e.g., Uber drivers) are contracted through new digital platforms and are assigned to work by intelligent algorithms (Ivanova et al., 2018; Jarrahi et al., 2020; Lee et al., 2018; Shapiro, 2018).

Against these backdrops, debates over autonomy issues, behavioural change, and AI have become increasingly visible in recent years (e.g., Griffiths & West, 2015; Kamphorst & Kalis, 2015; Loewenstein et al., 2015; Vugts et al., 2020). However, what is missing in the literature is a more nuanced approach to personal autonomy that examines specific situations and conditions where people's autonomy and agentic experience are influenced by different sources (e.g., other people, institutions, AI algorithms). Knowledge at this more nuanced level is especially important for developing more concrete and practical guidelines for designing behaviour change interventions as well as AI-infused systems that affect human behaviour. Building on previous theories and empirical works, we propose a functional model of personal autonomy to address this gap and examine specific autonomy-restricting situations and sources in three pre-registered experiments.

1.1 | Theoretical conceptualisations of personal autonomy

Autonomy is a multifaceted concept across different fields of inquiry. Traditionally, the concept is strongly rooted in the philosophy of human conduct. Early empiricist philosophers, such as Jeremy Bentham (1789) and John Stuart Mill (1861), represented autonomy in terms of unqualified freedom of choice. For Mill, autonomy entails complete individual liberty that should be restricted only by the requirement of protecting others. Modern scholars in philosophy, law and ethics generally agree that personal autonomy refers to an individual's capacity for self-governance. Beyond this, however, there is significant debate. Autonomy may include strong notions of doing the right thing, such as in neo-Kantian approaches to moral autonomy (cf., Korsgaard, 2009). Autonomy may also be represented as a more neutral condition where one is able to decide for oneself and pursue a course of action in one's life regardless of any particular moral content, generally labelled as procedural autonomy (Dworkin, 1988; Frankfurt, 1988).

The procedural view on autonomy resonates well with the Self-Determination Theory which conceptualises autonomy as being able to act in line with one's goals and experiencing volition over the goals that one enacts (Deci & Ryan, 2000; Ryan & Deci, 2000). Human behaviours are categorised on a continuum from fully autonomous to fully non-autonomous based on the types of motivations underlying the behaviours. Intrinsically motivated behaviours are by definition autonomous, while extrinsically motivated behaviours can be more or less autonomous depending on the degree of internalisations of the external regulation (Deci & Ryan, 2000). Moreover, autonomy is also considered an innate psychological need, and when supported by the environment, it promotes personal achievements and well-being (Deci & Ryan, 2000). The Self-Determination Theory has contributed to the empirical demonstration of the importance of autonomy in human functioning (Weinstein et al., 2012) and the design of interventions that support or even enhance autonomy (Ryan & Deci, 2019).

Finally, in the more applied area of bioethics, autonomy has been conceptualised as the competence of making unstrained decisions. This conceptualisation led to several frameworks that link personal autonomy to different behaviour change interventions. For example, the 'intervention ladder' categorises behaviour change interventions in terms of how much they restrict autonomy (Nuffield Bioethics Council, 2007). Restriction on autonomy increases as an intervention strategy moves from observation, education, to the change of choice architecture and incentives, and eventually to the elimination of choices. Griffiths and West (2015) revised the intervention ladder to redefine some intervention types as 'autonomy-enhancing', such as providing information and enabling choice. This modification was based on the idea that people are not always able to adhere to their goals so external forces may support autonomy in the sense of rational self-control (Walker, 2008). A similar line of reasoning pertains to recent advances in research on cognitive enhancement (e.g., Bostrom & Sandberg, 2009) and boosting (e.g., Grüne-Yanoff & Hertwig, 2016), pointing to the importance of external interventions that target mental functioning to increase personal autonomy in the service of goal achievement.

1.2 | A functional model of autonomy: The what, when, and how of goal-pursuit

So far, personal autonomy has been considered as a normative fact, a need construct, and a unidimensional construct mapped onto behaviour change interventions. Although informative, the previous conceptualisations do not distinguish different types of autonomy restrictions and how such restrictions impinge on agency, that is, the experience of control and authorship of actions (Aarts et al., 2005; Moore, 2016). To address these limitations, we propose a functional model of personal autonomy that considers the individual's opportunities to choose among behavioural options pertaining to a goal-pursuit in social and organisational contexts (e.g., going on a vacation or making a career plan). Our model builds on psychological and neurocognitive models of goal-directed behaviour, according to which agents are considered to be challenged by three main decisions, the *what*, *when* and *how* (Aarts & Elliot, 2012; Brass & Haggard, 2008). The *what* decision sets the goal a person is going to achieve in a given situation. In other words, it is the target behaviour someone decides to do in order to solve a problem, meet a challenge, or satisfy a need. The *when* decision determines the timing of act to perform the target behaviour (i.e., achieve the goal). The *how* decision defines the means through which the behavioural goal is achieved, since there are usually several means to achieve the same goal and the *how* decision selects one of the means or methods. In the same context, a *how* decision is subordinate to the corresponding *what* decision (see Carver & Scheier, 1982; Kruglanski, 1996) and multiple *how* decisions may be associated with the *same what* decision. For example, after setting a goal of having a summer vacation in Italy (the *what* decision), the decision-maker needs to decide the means of transportation, the duration of the trip, and the people who will join the trip (three *how* decisions).

The three decisions and their restrictions can be further illustrated by an example of curbing societal problems through behavioural interventions. When a government aims to change the way people commute to work in order to reduce pollution and/or congestion, they can either target the means of commuting (e.g., banning car-use and incentivising public transport; the *what* component), the timing of individual trips (e.g., restrictions on peak hours; the *when* component) and the execution of concrete actions given a travel mode (e.g., taking route A instead of B; the *how* component). These distinctive restrictions can also be understood in an organisational example of 'algorithmic management' (Kellogg et al., 2020; Schildt, 2017). A Uber driver's behaviours at work may be restricted in terms of which customer to pick up (the *what* component), when to pick up the customer (the *when* component), and through which route should they drive to the customer (the *how* component).

Extending these ideas to personal autonomy, one can consider autonomy as a direct function of the opportunities to decide what to do, when to act and how to act. Accordingly, removing any of the three components from one's control undermine autonomy, and potentially reduce experienced agency and motivation to pursue the personal goal at hand (i.e., goal motivation). While this functional model is intuitive and is grounded in psychology and neuroscience, it is yet to be tested in the context of social and organisational behaviours. For example, we do not know whether the three components contribute to personal autonomy independently and to the same extent. An empirically validated model can predict the effects of autonomy restrictions on human behaviour and inform intervention design in order to achieve desirable behaviour change without infringing too much autonomy.

1.3 | The nature of the restricting agent: Human versus AI

In traditional social and organisational contexts, a person's autonomy is usually restricted by another person, a small group of people, or all forms of social institutions (e.g., one's manager, a project team, or a government making policies). However, the recent development of AI technologies broadens the source of autonomy restriction from humans to machines that exhibit a certain degree of human intelligence. It has become increasingly common that people are told by AI agents—robots, chatbots or simply machine learning algorithms—rather than humans about what to do and when and how to do things (Sankaran et al., 2020). Therefore, a complete functional model of personal autonomy should consider whether people experience autonomy restrictions from AI agents differently than those from human agents and whether the relative importance of the *what*, *when* and *how* components shifts as a function of restricting agents.

There is a general negative sentiment towards the impacts of AI on human autonomy (Anderson et al., 2018). Empirical studies on autonomy per se are rare, but there is a tradition of research on people's trust and acceptance towards decisions made by AI agents versus humans. Classic studies by Dawes et al. demonstrated that in clinical contexts simple linear models often outperformed the intuitive judgments of clinicians, yet the clinicians showed strong reactance in using those models, a phenomenon they termed 'algorithm aversion' (Dawes, 1979; Dawes & Corrigan, 1974; Dawes et al., 1989). Recent studies have suggested a number of different causes for 'algorithm aversion'—AI agents were perceived to lack social and emotional intelligence for making complex or subjective decisions (Castelo et al., 2019; Lee, 2018), to be blind to people's personal preferences (Longoni et al., 2019), or were less tolerable when making errors (Dietvorst et al., 2015). However, the opposite phenomenon of 'algorithm appreciation' has also been reported—people sometimes prefer algorithmic to human judgments and decisions, especially when they lack competence themselves (Logg, 2017; Logg et al., 2019). In another study, participants reported a higher sense of autonomy when they were monitored by algorithms rather than by human managers in organisations, as they perceived machines to be less judgmental (Raveendhran & Fast, 2021). The debate on 'aversion and 'appreciation' aside, we ask an overlooked question—in addition to whether AI restricts more autonomy than humans in general, it is equally important to ask which types of decisions AI agents make and which types are perceived as more autonomy-restricting.

1.4 | The current investigation

The objectives of the current research are twofold. First, we set out to empirically test our functional model of personal

TABLE 1 Instructions about the source of autonomy restrictions in different conditions in Study 1a

Condition	Text in general introduction	Text in explaining the source of autonomy restrictions
Human	In this research, we are interested in how people respond to daily situations where the ways they act are partially determined by the decisions of another person. With another person, we do not mean a person who is part of your social life. The other person refers to someone who is neutral to you and is just part of the decision making context.	When the same decisions are made in the context of social interaction with another person, things become a bit more complicated. In such situations, it is not only you who decide 'what', 'when' and 'how' to do things. Often, the other person can also contribute to the interaction by making such decisions and thus determines the way you will act in the situation at hand.
AI	In this research, we are interested in how people respond to daily situations where the ways they act are partially determined by the decisions of an artificial intelligent (AI) agent. An AI agent is usually a computing system that exhibits certain intelligent capacities, such as perceiving the physical world, making decisions, and learning and adapting to its environments. In the present study, we do not mean any specific forms or applications of AI systems you may know of, but a decision-making agent powered by AI in general. The AI agent is neutral to you and is just part of the decision making context.	When the same decisions are made in the context of interaction with an AI agent, things become a bit more complicated. In such situations, it is not only you who decide 'what', 'when', and 'how' to do things. Often, the AI agent can also contribute to the interaction by making such decisions and thus determines the way you will act in the situation at hand.
Unspecified	In this research, we are interested in how people respond to daily situations where the ways they act are partially determined by constraints in these situations. Constraints in a situation can be physical, social or even personal (e.g., inability to decide). In the present study, we do not mean any specific kind of constraints, but focus on the mere fact that some aspects of decision-making can be constrained in the sense that a person is not able to decide.	When the same decisions are made in the context of constraints, things become a bit more complicated. In such situations, you may not be able to decide 'what', 'when' and 'how' to do things, but some of these aspects are constrained and these constraints determine the way you will act in the situation at hand.

autonomy in the context of social and organisational behaviours. We predicted that restricting any of the three components would undermine people's agency experience and their goal motivation in a decision-making situation. Without a-priori predictions, we also examine potential interactions between the components and the sizes of their relative impacts. Second, we aim to explore whether the same restrictions imposed by AI agents are perceived to be more negative than those imposed by humans and whether the source of restriction changes the weights of the three components.

Inspired by the policy capturing method (Aarts et al., 1997; Aiman-Smith et al., 2002), we conducted three scenario-based experiments where we systematically manipulated whether oneself or another agent had the opportunity to decide what to do, and when and how to do things.[1] In Studies 1a and 1b, we estimated the effects of the three components and source of restriction on experienced agency and compared their effect sizes in the generic context of personal goal-pursuit. In Study 2, we extended the paradigm to more concrete organisational settings and to test whether the same pattern of effects applied to people's goal motivation as reflected in their liking of a decision-making situation and their compliance with the situation. All three studies were reviewed and approved by the ethics review board of the Faculty of Social and Behavioural Sciences at Utrecht University.

[1] Data, analysis scripts, other materials, and the pre-registrations of the three experiments can be found in the Open Science Framework (OSF) repository: https://osf.io/65xhf/.

2 | STUDY 1A

2.1 | Method

2.1.1 | Participants

Four hundred and forty-four participants were recruited from Prolific, a popular online participant recruitment platform (https://www.prolific.co; see Peer et al., 2017 for an assessment of its data quality). The median age of the sample was 33.5 years, ranging between 18 and 64 years (SD = 12.4). There were 292 female participants. Most participants were from the United Kingdom or other English-speaking countries and they received 3.75 British pounds as compensation.

2.1.2 | Design

The experiment was a 2 × 2 × 2 (each goal-pursuit component: *what*, *when* and *how* determined by *oneself* or *another agent*) × 3 (source of restriction: *human*, *AI*, or *unspecified*) mixed design, with the three components manipulated within participants and source of restriction manipulated between participants. The unspecified source was added as a baseline condition to which the other two conditions were compared. Participants were randomly assigned to the human, AI or baseline conditions (144, 152 and 148 participants, respectively), where they received instructions that differed in certain parts that were about the source of autonomy

Block

Plan Your Holiday

IMAGINE THE FOLLOWING SITUATION: Holiday-season is approaching and you don't have a holiday plan yet. Therefore, you begin to plan your next holiday. In order to achieve your goal of having a holiday, three decisions need to be made:

- **What**: What will you do for your next holiday (e.g., *sightseeing in southern Italy* or *hiking in a nearby mountain range*);
- **When**: When to have your next holiday (e.g., *July* or *August*);
- **How**: How to realize the selected holiday (e.g., in case you travel to Italy: travel *by plane* or *by train*; in case you go hiking: hike for *one week* or *two weeks*).

You will be shown and respond to 8 scenarios where these decisions are made either by **yourself** or **another person**.

FIGURE 1 Vignette presented to the participants in the experimental block 'planning your holiday'

restrictions (see Table 1). For each participant, the experiment consisted of 32 trials, which were organised into four blocks concerning different types of goal-pursuit (*going on a holiday*, *making a work plan*, *improving one's health* and *arranging a social event*) presented in random order. Within each block, eight scenarios were randomly presented, where the determinations of the three components were manipulated.

2.1.3 | Sample size calculation

Based on a power analysis using the Superpower package in R (Lakens & Caldwell, 2021) and resource constrains, the initial plan was to recruit 450 participants. Because for a given sample size statistical tests for the within-subjects factors are much more powerful than for between-subjects factors or between-within interactions, the goal was to find a reasonable sample size that balance both requirements. The calculated sample size was more than enough for detecting small effects of the *what*, *when* and *how* components and their interactions as within-subject factors (assuming Cohen's $f = .15$), and also a small between-within interactions effect between each of the component and source of restriction with 73% power at the alpha level of .05.

2.1.4 | The scenario-based agency judgment task

The main experimental task was a scenario-based task of evaluating decision-making situations. At the beginning of each block, participants read a short vignette about a daily decision-making situation and were asked to imagine themselves being in the situation and making the decisions about the *what*, *when* and *how*. Depending on the condition, participants were told that one or more of the three components could be determined by another person (the human condition), an AI agent (the AI condition) or simply restricted (the baseline condition) (see Figure 1 for an example of the vignette provided to the participants[2]). Eight trials followed the short vignette and each trial presented the determinations of the three components through a diagram. For example, Figure 2a shows a scenario where all the components were determined by the participants themselves, while Figure 2b represents the case where the *when* and *how* components were determined by another person.

2.1.5 | Measurements

For each decision-making scenario, participants rated four items that are typical to assess the feeling of agency (e.g., Tapal et al., 2017): 'To what extent do you feel that you have freedom of choice in this scenario?', 'To what extent do you feel that you have control over the decision-making situation in this scenario?', 'To what extent do you feel that your autonomy is restricted in this scenario?' and 'To what extent do you feel responsible if you fail to achieve your goal in this scenario?'. Responses to the questions were given on a 7-point scale, ranging from 'not at all' (1) to 'very much' (7). Inter-item correlations were high between the four items (between .64 and .94), as well as the overall internal reliability (Cronbach's alpha = .93). We combined all four items into a single score of experienced agency.

2.1.6 | Procedure

Participants received invitations from Prolific to join the experiment implemented online on the Qualtrics website. First of all, participants read information about the experiment and clicked the 'Consent and Proceed' button to continue if they agreed to participate. Before the start of the main task, they were instructed about the meanings of the *what*, *when* and *how* components and that decisions regarding these components could be restricted by another agent.[3] They also practised the experimental task in two trials representing scenarios of *arranging a dinner*, a goal-pursuit type not used in the main task. Participants then continued to complete 32 trials in four randomised blocks with 30-s breaks in between. After the main task, participants answered some additional questions about the different types of goal-pursuit[4] and demographics. Finally, they were thanked, debriefed and redirected back to the Prolific website.

2.1.7 | Data analysis

We excluded some data based on two pre-defined criteria: (1) data from participants who read the critical instruction pages too fast (faster than 2.5 standard deviations from the means of the log-

[2] For brevity, other vignettes used in the three studies are not shown in detail in the manuscript but can be found in the OSF repository: https://osf.io/65xhf/.

[3] The definitions of the three decisions provided to the participants were very similar to those discussed in the Introduction. A complete description of the instructions can be found in the OSF repository: https://osf.io/65xhf/.

[4] These questions pertain to the relevance, difficulty and frequency of the goal-pursuit types in real life, but these variables had no impact on the main results.

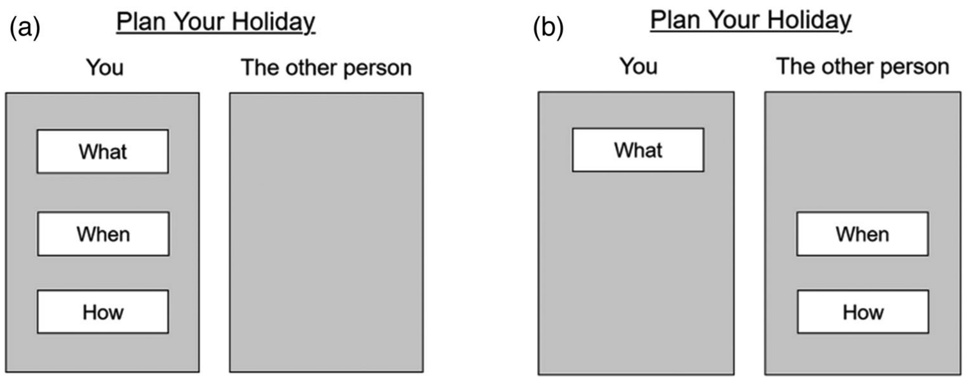

FIGURE 2 Two diagrams representing two exemplar scenarios of the goal-pursuit of planning a holiday

transformed page-reading time); (2) data from trials for which the concerned type of goal-pursuit was rated as irrelevant (relevance rating equalled 1). Applying these two criteria led to a remaining of 12,920 trials (90.9% of total trials) from 427 participants for statistical analyses.

To estimate the effects of the three components and their interactions on experienced agency and whether the effects were moderated by source of restriction, a linear mixed model was built for experienced agency with the four factors and their interactions as predictors, while allowing for random intercepts across the four types of goal-pursuit and across participants.[5] Deviation coding was used for the predictors so the regression coefficient for each main effect in the model could be interpreted as the average effect of a predictor (level one relative to the grant mean) across the different levels of other predictors. Since our large sample could reveal statistically significant but negligibly small effects, we also compared the percentage of variance accounted for between the full model and models with certain interaction terms removed, in order to interpret the effect sizes of the interactions. After obtaining the population-level regression weights for the three components, we went further to test the differences among them at the level of individual participants, by estimating person-specific effects of the three components from a random-slope model and then testing the differences between the person-specific weights in a linear-mixed model. All data analyses were performed using the R statistical software (version 4.03; R Core Team, 2020).

2.2 | Results

2.2.1 | Estimating the effects of the three components and their interactions

Figure 3 shows the mean of experienced agency as a function of levels of the three goal-pursuit components and the source of autonomy restriction. Evident by this visualisation, participants seemed to experience less and less agency when each of the three components was removed from their own control. This pattern was the same regardless of whether the source of autonomy restriction was another person, an AI agent, or unspecified. Given that the lines in Figure 3 were not completely parallel, interactions between the independent variables might exist, albeit very small in magnitude.

The effects of *what*, *when*, *how* and their interactions were estimated from linear mixed models for the three between-subjects conditions separately and were shown in Figure 4a. Confirming the visual pattern, a full linear mixed model revealed large main effects for all three components. Experienced agency was significantly and substantially higher when participants were able to decide the *what* (B_{what} = 0.98, 95% CI = [0.96, 0.99], p < 2e-16), *when* (B_{when} = 0.70. 95% CI = [0.69, 0.72], p < 2e-16) and *how* component (B_{how} = 0.83, 95% CI = [0.81, 0.84], p < 2e-16) in the imagined scenarios. The three main effects combined could account for over two-thirds of the variance in experienced agency (marginal-R^2 = 0.698).

There were also statistically significant interactions between the three components, but these effects were much smaller than the main effects (on the magnitude of 1/10), with the largest one being the three-way interaction ($B_{what*when*how}$ = 0.07, 95% CI = [0.06, 0.08], $p < .001$). As shown in Figure 3, the three-way interaction basically corresponded to a larger loss of experience agency when the number of goal-pursuit components determined by oneself changed from three to two and one to zero than when changed from two to one. All the interaction effects together only accounted for an extra 0.4% of the total variance in experienced agency.

While the effect estimates were largely consistent across the three between-subjects conditions (see Figure 4a), two small differences between the conditions were noticeable. First, the contribution of the *how* component in the human and AI conditions were significantly larger than in the baseline condition ($B_{how*human}$ = 0.07, 95% CI = 0.05, 0.09], p < .001; B_{how*AI} = 0.04, 95% CI = 0.02, 0.06], p < .001). Second, there was a small but significant main effect of condition AI ($B_{condition}$ = −0.10, 95% CI = [−0.16, −0.04], $p < .001$), indicating participants experienced less agency when interacting with an AI agent, regardless of the opportunity to determine the three components by oneself. Note that these small effects only account for an extra 0.5% of the total variance in experienced agency.

[5] We also tested whether the main results varied across the different types of goal-pursuit, but no noteworthy differences were found in both studies.

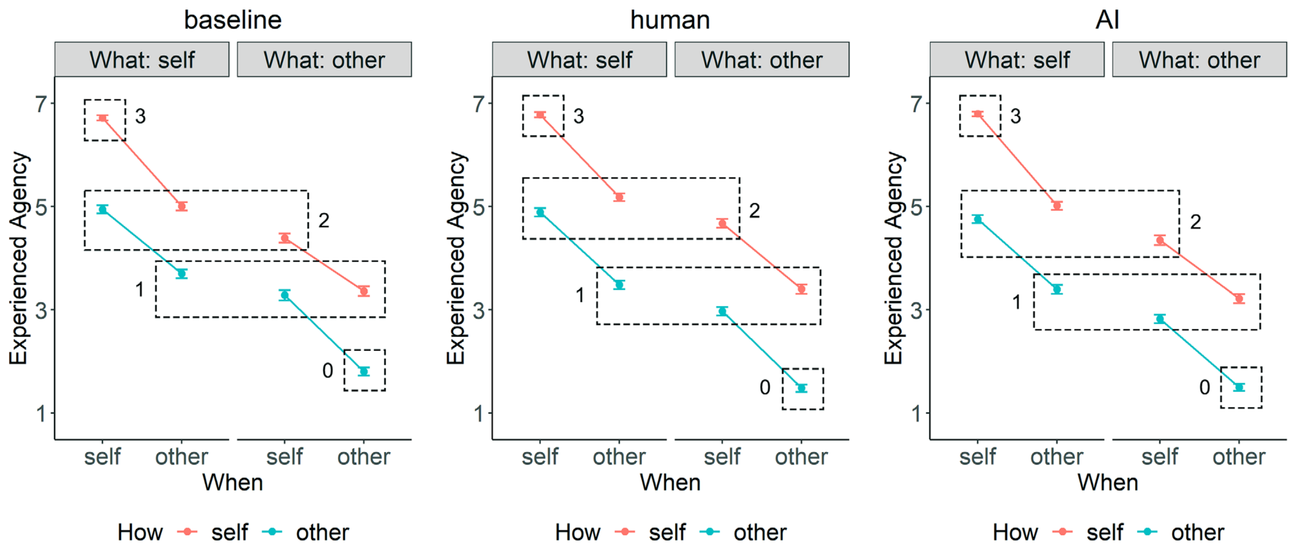

FIGURE 3 Experienced agency as a function of the levels of the three goal-pursuit components (self vs. other) and the source of autonomy restriction (baseline, human and AI) in Study 1a (error bars representing 95% confidence intervals). The means are also grouped into dashed boxes that indicate the number of goal-pursuit components determined by oneself.

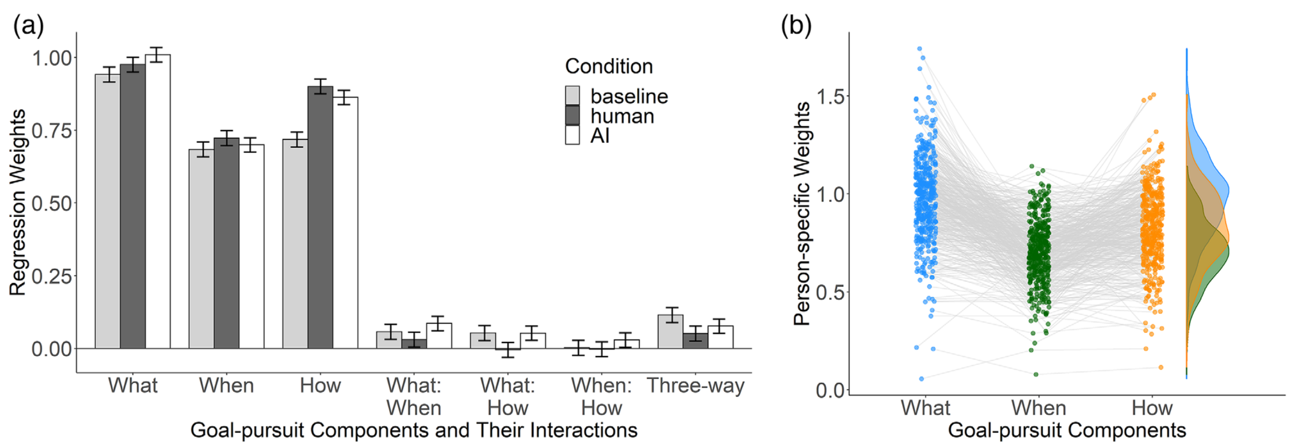

FIGURE 4 (a) Comparing the regression weights (*B*) of the three goal-pursuit components across the three between-subjects conditions in Study 1a (Error bars representing 95% CIs); (b) Distributions of person-specific weights of the three components and the within-person patterns in Study 1a.

2.2.2 | Comparing the effects of the three components at the individual level

Figure 4b shows the person-specific effects of the three components on experienced agency. The linear-mixed model revealed that the *what* component was weighted more than the *how* component (*B*= 0.15, 95% CI = 0.13, 0.17], $p < .001$), and the *how* component was weighted more than *when* component (*B*= 0.12, 95% CI = [0.10, 0.15], $p < .001$). The differences among the three components accounted for 24.5% of the variance in the person-specific regression weights, while individual differences could explain an additional 19.7%. As indicated by the crossing lines in Figure 4b, while the ranking of regression weights applied to the majority of participants, it was nonetheless drastically different for some participants, for example, for whom the when component influenced experienced agency the most.

2.3 | Discussion

Study 1a supported the idea that all three components were important for personal autonomy, as removing them from one's control undermined experienced agency. Their relative impact followed a clear linear ranking: the *what* component had the largest impact, followed by *how* and finally *when*. As indicated by the negligibly small interaction effects, the effects of the three components were additive, as none of the components had an overriding influence on experienced agency alone.

Furthermore, the ranking of importance was the same, regardless of the nature of the constraining agent (human, AI or unspecified).

The small interaction effects may reflect more about how participants used the rating scale in the experimental task, rather than anything theoretically interesting about autonomy. In the two extreme scenarios (all three components or no component determined by oneself), participants might have simply chosen the two extremes (one and seven). When one or two components were determined by oneself, participants might tend to distance their ratings away from the two extremes. This hypothetical strategy would generate the unequal decreases in experienced agency among different pairs of scenarios.[6]

3 | STUDY 1B

In Study 1a, the verbal explanation and visual representation of the components always followed the same order of *what*, *when* and *how*, which might have contributed to the estimated importance of the components. For example, the regression weight of the *what* component could have been inflated because participants always saw it at the top of the diagrams. In Study 1b, we aimed to replicate Study 1a, while ruling out this confounding factor. Because source of restriction did not matter much in Study 1a, only the human agent condition was included in Study 1b.

While Study 1a showed the value of the paradigm to reveal the relative importance of the components, one might ask whether the same weights could be obtained through a simpler method of directly asking people. Previous research using the policy-capturing method suggested the answer is 'no' since regression weights estimated statistically often differ substantially from subjective importance explicitly rated by participants (Barlas, 2003; Brookhouse et al., 1986; German et al., 2016; Schmitt & Levine, 1977; Slaughter et al., 2006). To explore this issue, in Study 1b we asked participants to directly rate the importance of *what*, *when* and *how* components for the different types of goal-pursuit and compared the subjective importance with the regression weights estimated from the experimental task.

3.1 | Method

3.1.1 | Participants

One hundred and twenty participants were recruited from Prolific. The median age was 32 years, ranging between 18 and 64 years old (SD = 11.3). Similar to Study 1a, there were more females (77) than males (42)and 1 "other" (non-binary) and most participants were from the United Kingdom or other English-speaking countries. During the recruitment, two participants were rejected because they answered all questions after each scenario with the same ratings, even for the reverse-coded item. All accepted participants received 2.75 British pounds as compensation.

3.1.2 | Design

The basic design of the experiment was a replication of the human condition in Study 1a. In addition, we counterbalanced the order of explaining and presenting the three components: participants were randomly assigned to one of the three conditions with different orders, namely 'what-when-how', 'when-how-what' and 'how-what-when'. For each participant, the block and trial structures were identical to Study 1a.

3.1.3 | Sample size estimation

Based on the effect estimates from Study 1a, a power analysis using the R package Superpower resulted in a sample of 120 participants, in order to have 90% or higher power (alpha = .05) to (1) replicate the same effect sizes as with the same order condition in Study 1a; (2) detect a weaker overall effect size for the differences between estimated effects (mean difference = .085); and (3) reveal small between-within interaction effects between the three components and the order of presenting the components.

3.1.4 | The scenario-based agency judgment task

The same task was used in Study 1b, except that the order of explaining and visually presenting the components was varied according to the order condition assigned to each participant. Figure 5 shows the same two scenarios from Figure 2a but with different presenting orders for the 'when-how-what' (Figure 5a) and 'how-what-when' condition (Figure 5b).

3.1.5 | Measurements

The same items were used to measure experienced agency. Similar to Study 1a, inter-item correlations were high between items (between .59 and .95), as well as overall internal reliability (Cronbach's alpha = .93). Thus, the four items were combined into a single measure of experienced agency for statistical analyses. At the very end of the experiment, participants were also asked to rate the importance of the *what*, *when* and *how* components to each of the goal-pursuit types on a scale from 1 (not at all important) to 10 (extremely important).

3.1.6 | Procedure

The study followed the same procedure as with Study 1a. The only change was that in order to enforce participants to read important

[6] Given that the interaction effects were negligibly small and theoretically uninteresting, we will only report estimated effects (e.g., Figure 4a) but not interaction patterns (e.g., Figure 3) in the rest of the article. Interested readers can find more results and visualisations in the Online Supplemental Material.

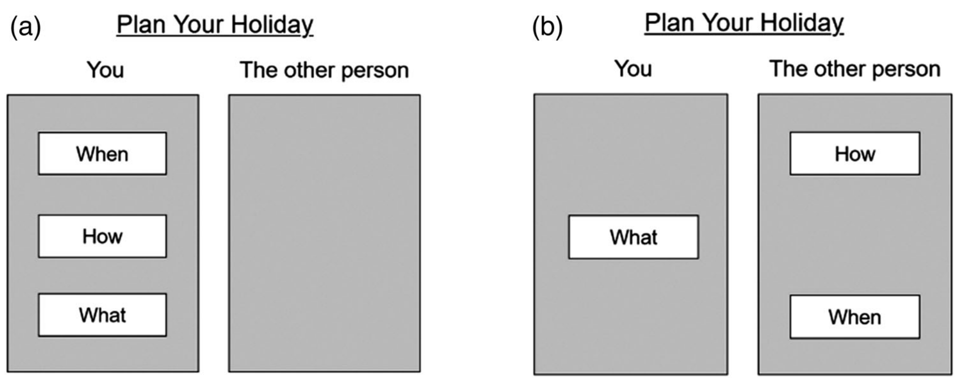

FIGURE 5 Two diagrams representing the same two scenarios in Figure 2 but with different orders of the three components

instruction pages carefully, timers were used so that only after a certain time had passed (e.g., 10 s) could participants click the button to proceed to the next page. The new strategy was adopted because in Study 1a that quite a lot of participants (9%) skipped through these pages and were rejected.

3.1.7 | Data analysis

Applying the same two criteria as in Study 1a led to 3664 valid trials (95.4% of total) from 119 participants for statistical analysis. We used the same linear mixed modelling approach to estimate the effects of the three components, the order condition, and interaction between these factors, with random intercepts modelled across blocks and participants. As with Study 1a, we also tested the differences among the person-specific regression weights estimated from a random-slope model. Finally, we analysed the subjective rating of importance, testing whether the same linear ranking would emerge and to what extent they correlated with the statistically estimated importance from the experimental task at the individual level.

3.2 | Results

3.2.1 | Replicating Study 1a with counter-balanced order

Figure 6a shows the effect estimates of the three components and their interactions for the three order conditions separately. The full linear mixed model revealed three strong main effects—experienced agency was judged to be much higher when participants were able to decide the *what* (B_{what} = 0.98, 95% CI = [0.95, 1.01], p < 2e-16), *when* (B_{when} = 0.71. 95% CI = [0.68, 0.74], p < 2e-16) and *how* (B_{how} = 0.79, 95% CI = [0.76, 0.82], p < 2e-16) in the study scenarios (nearly 70% of the variance in experienced agency was accounted by the three components). The only significant interaction effect relating to the three components was the three-way interaction effect ($B_{what*when*how}$ = 0.09, 95% CI = [0.06, 0.012], p < .001). The interaction effects only accounted for an extra 0.4% of the variance in experience agency.

More importantly, confirming the consistent pattern across the order conditions in Figure 6a, only two negligibly small moderation effects of the order were noteworthy. Compared with the average effect of the *what* component on experience agency, its influence was slightly smaller in the 'what-when-how' condition (B = −0.07, 95% CI = [−0.11, −0.04], p < .001) but slightly larger in the 'when-how-what' condition (B = 0.08, 95% CI = [0.04, 0.12], p < .001). Although these effects were statistically significant, they only accounted for an extra 0.3% of the total variance in the outcome variable. Also, the directions of the effects were opposite to what the 'order as potential bias' hypothesis would predict.

At the level of individual participants, person-specific weights decreased from *what* to *how* (B = −0.19, 95% CI = [−0.22, −0.15], p < 2.2e-16) and from *how* to *when* (B = −0.08, 95% CI = [−0.12, −0.04], p < .001) (see Figure 6b). Indicated by marginal-R^2, the differences across the three components accounted for 28.4% of the variance in the person-specific regression weights. Despite the overall reliable ranking at the group level, some participants did exhibit deviating patterns, for example, for whom experienced agency was influenced the most by the *when* component. These individual differences accounted for an additional 21.5% of variance in the person-specific regression weights.

3.2.2 | Comparing subjective rating of importance and statistically estimated importance

Figure 7a shows the average subjective importance of the three goal-pursuit components for different types of goal-pursuits rated by the participants. The subjective importance of all three components was much lower when the goal-pursuit was about arranging social events, even though this difference did not emerge in estimated regression weights. The linear-mixed model indicated that the differences among the subjective importance of the three components were very small (M_{what} = 7.72, SD_{what}= 1.18; M_{when} = 7.29, SD_{when} = 1.51; M_{how} = 7.46, SD_{how} = 1.33) and they accounted for only 1.7% of the total variance in subjective importance. In contrast, there were large individual

FIGURE 6 (a) Comparing the regression weights (*B*) of the three goal-pursuit components across the three order conditions in Study 1b (Error bars representing 95% CIs); (b) Distributions of person-specific weights of the three components and the within-person patterns in Study 1b.

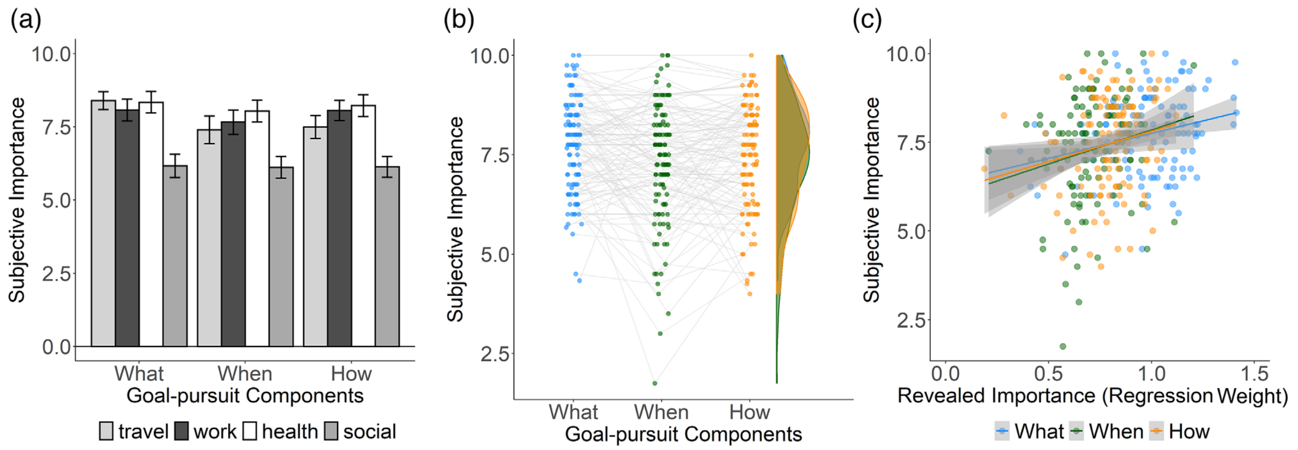

FIGURE 7 (a) Subjective importance for the three components and different types of goal-pursuits (error bars representing 95% CIs); (b) Distributions of subjective importance and the within-person patterns; (V) Correlations between subjective importance and statistically estimated importance for the three goal-pursuit components.

differences in terms of which of the components were considered to be subjectively more important, accounting for about 50% of the total variance in the subjective rating (see Figure 7b). Finally, subjective importance of the three components did not correspond closely to the statistically estimated importance, as evidenced by the rather weak correlations between them (*what*: $r = 0.26$, $p = .004$; *when*: $r = 0.19$, $p = .038$; *how*: $r = 0.20$, $p = .033$; see also Figure 7c). This was corroborated by the fact that only 35 participants (29.4%) reported the same ranking of importance as the estimations from their experimental data.

3.3 | Discussion

Study 1b successfully replicated the findings in Study 1a and ruled out the possibility that the explanation and presentation order of the three components contributed to the ranking of their effect sizes on experienced agency. The results also suggest that the differences in relative importance of the *what*, *when* and *how* components are not detectable by simply asking people to rate their importance on Likert scales.

4 | STUDY 2

Studies 1a and 1b tested the functional model of personal autonomy in generic personal goal-pursuit context (e.g., planning a holiday, arranging a social event). The results suggest robust main effects of the three components on experienced agency and a clear ranking of their relative importance. The interaction effects between the components and the differences between human and AI as autonomy-restricting agents were negligible. In Study 2, we aimed to extend the paradigm to the organisational context and more concrete settings (e.g., signing up for a job, making a career development plan). Instead of referring to the autonomy-restricting agent in a generic way as 'another person' or 'an AI agent', more details and contextual information were

TABLE 2 Instructions about the source of autonomy restrictions in different conditions in Study 2

Condition	Text in explaining the source of autonomy restrictions
Human	In organisational contexts, the three aspects above (the what, when and how) are often not completely up to yourself. Rather, one or more are decided by your manager for you. When making these decisions, your manager may use various sources of information to find options that satisfy the needs of the organisation but also account for your own interests. In this study, we are interested in the general question of how you respond to different situations where the three decisions are allocated to yourself and your manager differently.
AI	In recent years, with the rapid development of artificial intelligence (AI) technologies, it has become more and more common that organisations use AI algorithms to manage their employees. Thus, in organisational contexts, the three aspects above (the what, when and how) are often not completely up to yourself, but are decided by an AI algorithm for you. When making these decisions, the AI algorithm may use data from various sources to find options that satisfy the needs of the organisation but also account for your own interests.

given about the identities of the agents and their relationships with the decision-maker,that is, a human manager or an AI algorithm used by an organisation to manage its human resources.

One limitation of Studies 1a and 1b was the measurement of experienced agency as our only dependent variable. While agency experience closely relates to autonomy and is important for human functioning (Aarts et al., 2005; Moore, 2016), organisational management or interventions usually concern more consequential variables such as employees' motivation to pursue their goals given the autonomy restrictions in the organisation (e.g., to participate in a training or health promotion programme). Study 2 tested whether and how restrictions on the three components reduce goal motivation by measuring participants' likings of a decision-making situation and their intentions to comply with the situation.

4.1 | Method

4.1.1 | Participants

One hundred participants were recruited from Prolific (87 women and 13 men). Given the organisational context, we restricted our recruitment for participants living in the United Kingdom and employed at the time of the study. The median age was 31 years, ranging between 20 and 54 years old. During the recruitment, two participants were removed as they repeatedly answered all questions after each scenario with the same ratings, even for the reverse-coded items. All accepted participants received three British pounds as compensation.

4.1.2 | Design

Study 2 followed a 2 × 2 × 2 (each goal-pursuit component: *what*, *when* and *how* determined by *oneself* or *another agent*) × 2 (source of restriction: *human manager*, *AI algorithm*) mixed design. Participants were randomly assigned to the human or the AI conditions (see Table 2 for the different texts used to explain the two different sources of autonomy restrictions). Each participant went through four types of goal-pursuit in an organisational context—*assigning food-delivery tasks*, *making a career development plan*, *participating in an occupational health intervention programme*, and *joining a personal training programme*. For each goal-pursuit type, eight scenarios were randomly presented, where the determinations of the three components were manipulated.

4.1.3 | Sample size estimation

Based on the effect estimates from Study 1a, a power analysis resulted in a sample of 100 participants, in order to have 90% or higher power (alpha = .05) for the within-subjects effects: (1) the same effects of autonomy restrictions on experienced agency as in Study 1a; (2) slightly weaker effects of autonomy restrictions on liking and intention. The same sample size could also detect a medium between-subjects effect of source of restriction (human manager vs. AI algorithm; Cohen's $f = .29$) with a power of 81.6% at the alpha level of .05.

4.1.4 | The scenario-based agency judgment task

Through the same procedure of reading vignettes as in Studies 1a and 1b, participants were asked to imagine themselves being in an organisational decision-making situation (e.g., assigning food-delivery tasks) and consider the *what*, *when* and *how* components (e.g., which customer to serve or which order to pick-up, when is the assigned delivery, and how to deliver the food to the destination). Depending on the condition, participants were told that these decisions could also be made by their manager in the organisation or an AI algorithm used by the organisation to manage its employees. They were also told that the human manager (or the AI algorithm) would consider the interests of the organisation and the employee (i.e., the participant in the hypothetical scenario) by processing data and information from various resources (e.g., in the case of assigning the food-delivery task, the delivery history of the employee, current requests from customers and the travel location, etc.).

In each scenario, the determinations of three components were visualised using the same diagram as in Study 1a. Participants were asked to report experienced agency, liking of the decision-making situation and intention to comply with the situation (e.g., whether or not to sign up for the food-delivery job after learning about how the different decisions were made).

4.1.5 | Measurements

The same four items were used to measure experienced agency and were combined into a single measure for statistical analyses (Cronbach's alpha = .95). We only slightly rephrased the responsibility question to be 'To what extent do you feel responsible if the decisions made in this scenario turn to be bad?' in order to adapt to the organisational context. Liking was measured by two items—'Assuming you are facing the decision-making situation in this scenario in your real life, how much would you like it?' (liking) and 'Assuming you are facing the decision-making situation in this scenario in your real life, how much would it annoy you?' (annoyance)—using 7-point response scales (1 = not at all; 7 = very much). The two items were combined into a single measure of liking with very high internal reliability (Cronbach's alpha = .93). Participants' intentions to comply with the situation were measured by a single question; for example, in the case of assigning food-delivery task as 'Assuming you are facing the decision-making situation in this scenario in your real life, to what extent would you intend to sign up for the food-delivery job?'. The same 7-point scale was used.

4.1.6 | Procedure

The study followed the same general procedure as with the two previous studies.

4.1.7 | Data analysis

We exclude the trials (2.1%) where participants responded to the reverse-coded items for experienced agency and liking in the same way as with the corresponding items (e.g., someone answered 'not at all' for both the liking and annoyance questions). After data cleaning, we used the same statistical models and procedures as in Study 1a to test the influence of the three components as well as source of restriction (human manager vs. AI algorithm) on experienced agency, liking and intention to comply with a decision-making situation.

4.2 | Results

4.2.1 | Replicating the results for experienced agency

Figurer 8a shows the effect of *what*, *when*, *how* and their interactions on experienced agency from linear mixed models for the two agent-type conditions separately. Being able to decide on *what* (B_{what} = 0.96, 95% CI = [0.93, 0.99], p < 2e-16), *when* (B_{when} = 0.90, 95% CI = [0.87, 0.93], p < 2e-16) and *how* (B_{how} = 0.82, 95% CI = [0.78, 0.85], p < 2e-16) component was associated with higher experienced agency (accounted for 73% of the variance in experienced agency in total). The only noticeable interaction effect was the positive three-way interaction effect ($B_{what*when*how}$= 0.06, 95% CI = [0.03, 0.09], p < .001), but again the effect was less than one-tenth of the main effects. We did not find a significant main effect of source of restriction: the scenarios with the human superordinate versus an AI algorithm did not lead to any difference in experienced agency (B_{agent} = −0.001, 95% CI = [−0.07, 0.07], p < .56). Source of restriction did have a small and unexpected moderating effect—the regression weight of *when* was significantly higher in the AI versus in the human condition ($B_{when*agent}$ = 0.05, 95% CI = [0.02, 0.08], p < .001), meaning that when being managed by an AI algorithm, participants' agency experience depended more heavily on the determination of *when* to pursue one's goals.

Instead of the ranking of *what*, *how* and *when* found in Study 1a and 1b, the estimated person-specific weights suggested a ranking of *what*, *when* and *how* (see Figure 8b). A linear mixed model indicated that 10.4% of the variance in person-specific weight was accounted for by the different components, and 27.1 accounted by individual differences. The *what* component influenced experienced agency more than the *when* component (B= 0.06, 95% CI = [0.02, 0.10], p = .004), and the *when* component influenced experienced agency more than the *how* component (B= 0.09, 95% CI = [0.05, 0.13], p < .001). In addition, the scenario of assigning food-delivery tasks clearly resulted in a different ranking of *when*, *how* and *what*. (see Figure 8c). Adding goal-pursuit type as a predictor in the model showed that for assigning good-delivery tasks, *when* became more influential (B= −0.18, 95% CI = [−0.27, −0.10], p < .001) and *what* became less influential (B= 0.22, 95% CI = [0.13, 0.30], p < .001).

4.2.2 | Extending the analyses to liking and behavioural intention

Overall, restrictions on the three components influenced participants' liking and behavioural intentions in the same way as they influenced experienced agency (see Figures 9a and 10a). Restricting the *what* (B_{what} = 0.70, 95% CI = [0.65, 0.74], p < 2e-16), *when* (B_{when} = 0.73, 95% CI = [0.69, 0.77], p < 2e-16) and *how* component (B_{what} = 0.65, 95% CI = [0.60, 0.69], p < 2e−16) made participants judge decision-making situations to be more negative (marginal R^2 = 0.469). In contrast to the large main effects, no interaction effects on liking were larger than 0.1. Participants showed no differences between restrictions from a human manager and an AI algorithm in terms of liking or disliking a decision-making situation (B_{agent} = −0.04, 95% CI = [−0.13, 0.06], p = .46), nor did source of restriction moderate the effects of the three components (all *ps* > .12). Finally, when person-specific weights were considered, the differences across the components accounted for only 1.7% of the total variance in liking, while individual differences accounted for 67.3% of the total variance.

For intention to comply with a decision-making situation, we again found strong and significant main effects of the three components. When participants were not able to decide the *what* (B_{what} = 0.63, 95% CI = [0.59, 0.67], p < 2e-16), *when* (B_{when} = 0.66, 95% CI = [0.62, 0.70], p < 2e-16) and *how* (B_{how} = 0.57, 95% CI = [0.53, 0.62], p < 2e-16) of goal-pursuit by themselves, their intention to go along with a

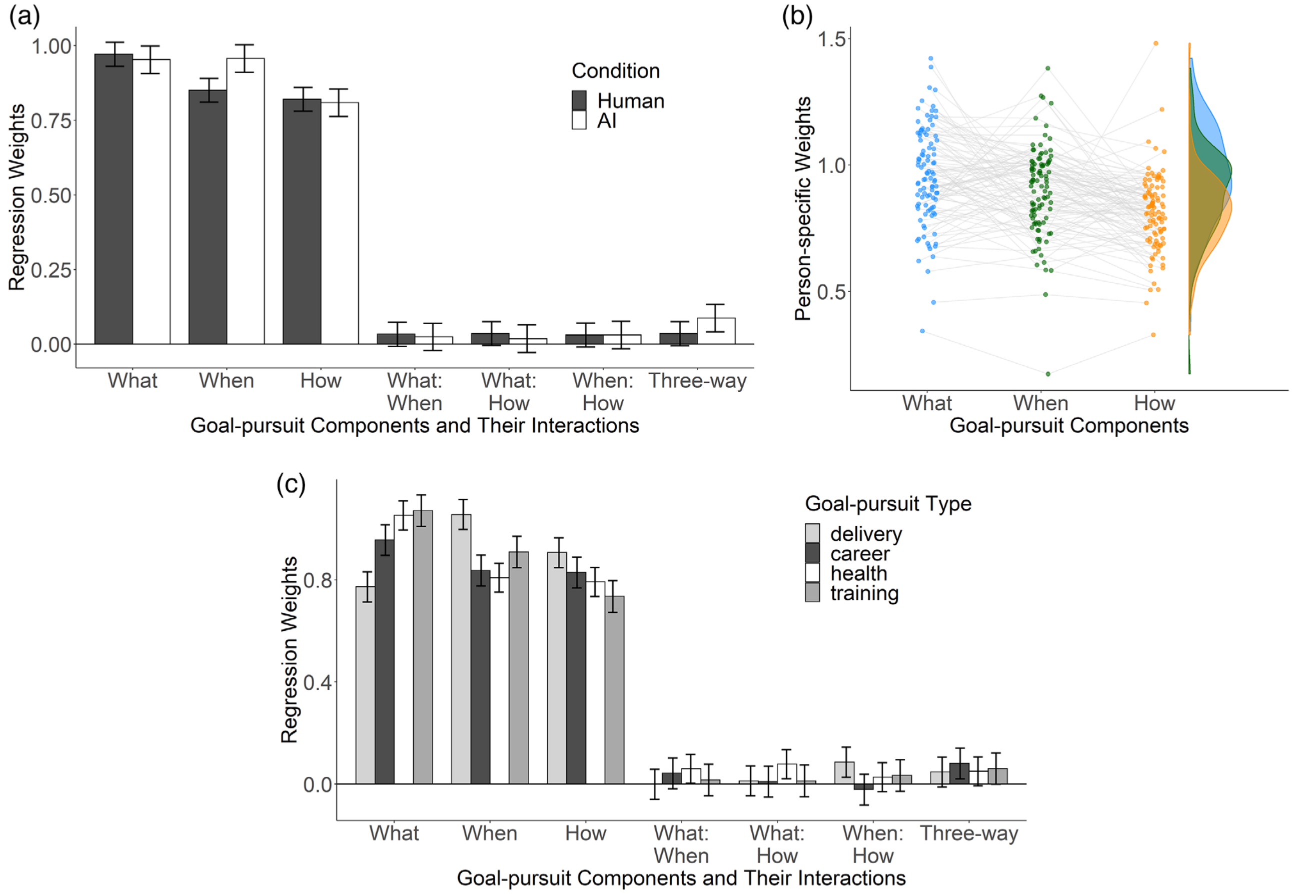

FIGURE 8 (a) Regression weights (*B*) of the three components on experienced agency across the two between-subjects conditions (Error bars representing 95% CIs); (b) Distributions of person-specific weights of the three components on experienced agency and the within-person patterns in Study 2; (c) Regression weights (*B*) of the components across different types of goal-pursuit (delivery = assigning a food-delivery task, career = making a career plan, health = designing a health intervention program, training = joining a personal training).

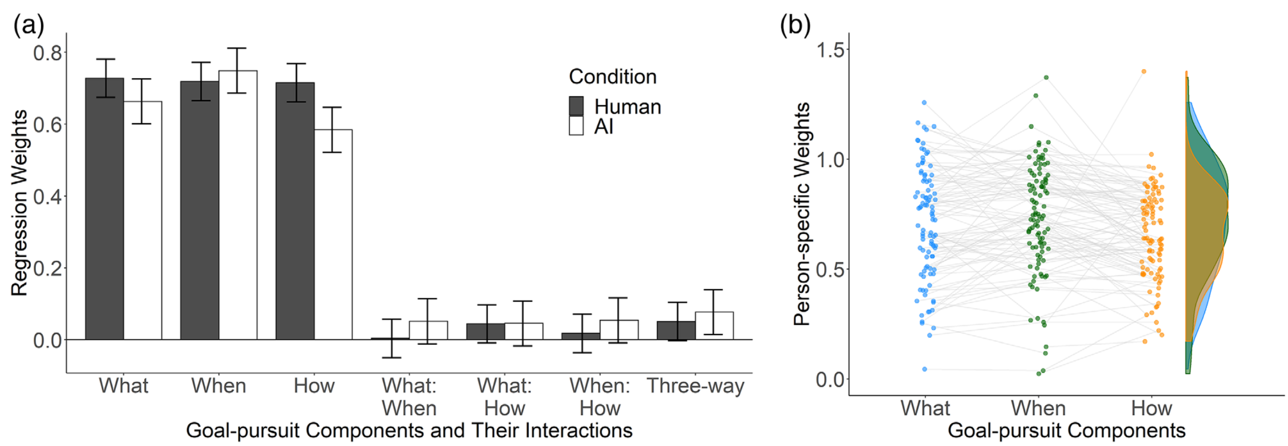

FIGURE 9 (a) Regression weights (*B*) of the three goal-pursuit components on liking across the two between-subjects conditions in Study 2 (Error bars representing 95% CIs); (b) Distributions of person-specific weights of the three goal-pursuit components on liking and the within-person patterns in Study 2.

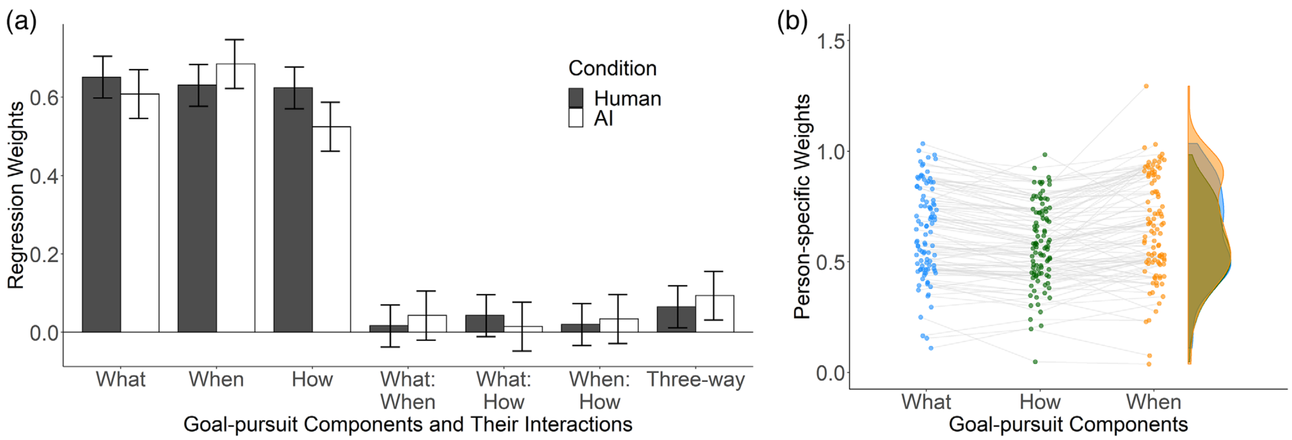

FIGURE 10 (a) Regression weights (*B*) of the three goal-pursuit components on compliance intention across the two between-subjects conditions in Study 2 (Error bars representing 95% CIs); (b) Distributions of person-specific weights of the three goal-pursuit components on behavioural intention and the within-person patterns in Study 2.

decision-making situation became much weaker (marginal $R^2 = 0.362$). Again, no interaction effects were larger than 0.1. Source of restriction did not influence intention ($B_{agent} = -0.05$, 95% CI = [−0.18, 0.08], $p = .45$), nor did it contribute to any interaction effects. Finally, the differences across the components accounted for only 2.1% of the total variance in liking, while individual differences accounted for 85.7% of the total variance.

4.3 | Discussion

Study 2 partially replicated the results of Studies 1a and 1b. The strong main effects of restricting the *what*, *when* and *how* component and the negligible interaction effects between them extended from experienced agency to goal motivation as reflected in liking and compliance intention. However, the ranking of *what*, *how* and *when* in their relative influences on experienced agency was not found, and the differences among the person-specific weights of the components were much smaller. In terms of our functional model of autonomy, it did not matter if the restricting agent was a human manager or an AI algorithm used by an organisation.

5 | GENERAL DISCUSSION

The present study proposed a functional model of personal autonomy and tested the model in three scenario-based experiments. Studies 1a and 1b suggest that in general situations of personal goal-pursuit, restricting any of the *what*, *when* and *how* components reduces experienced agency but the interactions between the components are negligible. The two studies also indicate the importance of the three components to follow a ranking of *what*, *how* and *when*. Study 2 generalises the effects of goal motivation in an organisational context and shows that the ranking of importance might be more context-dependent. No study suggests the functional model to be sensitive to the nature of the restricting agent (human vs. AI).

5.1 | Implications for theories of personal autonomy

With a strong root in moral philosophy, the concept of personal autonomy has been incorporated into several influential psychological theories about social motives, including the Self-Determination Theory (Deci & Ryan, 2000; Ryan & Deci, 2000), theories of psychological reactance (Brehm, 1966; Rosenberg & Siegel, 2018), and theories about the need for control and choice (Leotti et al., 2010). Research driven by these theories often focuses on the consequences of autonomy deprivation in terms of general human functioning (e.g., Weinstein et al., 2012) or specific behavioural responses that are motivated to restore freedon of choice and control (e.g., Rosenberg & Siegel, 2018). At the practical side, debates are centred on whether specific behavioural change interventions or autonomous technologies undermine personal autonomy (e.g., Griffiths & West, 2015; Kamphorst & Kalis, 2015; Loewenstein et al., 2015; Väänänen et al., 2021; Vugts et al., 2020). Contributing to this broad literature, our work connects the theoretical and practical sides through a functional model that posits the degree of being autonomous as a function of the opportunities to decide *what*, *when* and *how* in social contexts of goal-pursuit. Supporting this model, our studies suggest that all three components independently and substantially contribute to personal autonomy. None of the three components overrides the influences of the other components, indicated by the lack of interaction effects.

The ranking of *what*, *how* and *when* in terms of their importance found in Studies 1a and 1b is consistent with a large body of literature on human decision-making and behaviour. For example, the *what-how* distinction is similar to the conceptualisation of means-end relationships in psychological theories (Carver & Scheier, 1982; Kruglanski, 1996). Deciding *what to do* in the context of goal-pursuit is at a superordinate level while deciding on the specifics as to *how to achieve the goal* is at a subordinate level. Similarly, artificial agents who can decide on ends are also considered to be more autonomous than those who can only choose between means (e.g., Corchado et al., 1997; Luck & d'Inverno, 1995). Furthermore, the superiority of *what* to *when* resonates with the *what-when-whether* model of intentional action in the neuroscience of willed action (Brass & Haggard, 2008). Neurophysiological evidence suggests that *what* and *when* decisions are dissociable in terms of their underlying brain regions, and that *what* decisions have a stronger contribution to voluntary action execution compared to *when* decisions (e.g., Hoffstaedter et al., 2013; Krieghoff et al., 2009; Zapparoli et al., 2018).

However, the idea that this ranking is universal was rejected by Study 2—the differences among the relative importance of the three components were much smaller in the organisational context and the ranking was context-dependent. One hypothesis for this discrepancy is that the stability of the ranking may depend on the abstractness of concreteness of the experimental scenarios. When the scenarios are highly abstract (e.g., simply referring to 'another person' or 'an AI agent'), people may find it difficult to evaluate the context-specific importance of the components and this relies heavily on a ranking of *what*, *how* and *when*, which is implied in language and culture. Once the scenarios are made more concrete, people become more capable of distinguishing the importance of the components to specific decision-making situations. Future research is required to test this hypothesis.

5.2 | Implications for the debate over human and algorithmic decision-making

Our findings also contribute to the literature on human versus algorithmic decision-making and the debate over whether people appreciate or are aversive to decisions made by AI systems (Castelo et al., 2019; Dietvorst et al., 2015; Lee, 2018; Longoni et al., 2019; Logg, 2017; Logg et al., 2019). Our results suggest that in terms of the functional model of personal autonomy, whether the autonomy restrictions come from a person or an AI algorithm does not matter much, or at least much less than the determinations of the specific goal-pursuit components. The second study further demonstrates that people's motivation to pursue goals do not become weaker when an AI algorithm instead of a human takes the role of management. Since all our experiments are scenario-based, whether the 'human-AI equality' holds in the real world remains a question for future research. Nonetheless, our research points to the importance of not treating AI as a general category but examining what AI systems actually do (e.g., restricting *what*, *when* or *how*).

5.3 | Implications for behaviour change in the organisational context and beyond

Our findings also have practical implications for policy-makers, employers and intelligent system designers who are interested in managing or changing human behaviours. First, the decomposition of goal-pursuit into three distinct components offers a new perspective on the design space of interventions. Traditionally, intervention designers may choose from a series of intervention techniques that are supposed to undermine personal autonomy to a different degree on a unidimensional scale (Nuffield Bioethics Council, 2007). Instead, our model reminds one that changing behaviours in the real world often implies restrictions for one or more of the *what*, *when* and *how* components. For example, in order to promote healthy diets, interventions may be directed at what people eat, when they eat and how (much) they eat. Intervention designers should also be aware of the specific goal-pursuit component for which the choice is eliminated or guided and carefully weigh the effectiveness of targeting a specific component and its potential negative impact on agency and goal motivation.

Second, the additive effects of the three components suggest a 'compensation' mechanism that can be utilised by intervention designers. The lack of intervention effects means that when one of the components is restricted, freeing another component can partially make up for the negative effects caused by the first component. For example, when a manager plans to exert more control over the behaviours of employees, a strategy might be to reduce their freedom in choosing what projects they wish to work on. This intervention will inevitably undermine the employees' experienced agency and may further harm intrinsic motivation (Deci & Ryan, 2000), but this adversity can be partially restored by allowing more flexibility in when to start and finish their projects. To make the 'compensation' effective, designers may follow the revised 'intervention ladder' (Griffiths & West, 2015) to choose intervention techniques that enhance autonomy to target one of the three components (e.g., providing information, enabling choice).

It should be noted that autonomy-enhancing interventions have also been put into practice through the lens of the Self-Determination Theory, most notably in application areas of education, organisational behaviour and technology design (Ryan & Deci, 2019). The practice often involves the enabling of choices relating to the three goal-pursuit components in our model, but the distinction among them has not been discussed formally. Thus, our model and empirical findings open up some interesting intervention research questions and design opportunities for enhancing autonomy from the perspective of Self-Determination Theory.

Third, we provide an experimental paradigm to access their weights in specific application situations. The method can be easily and quickly managed in companies and other organisations. Our results corroborate with previous literature to suggest the policy-capturing method as a potentially more sensitive and less biased alternative to self-report measures to reveal such information (Barlas, 2003; Brookhouse et al., 1986; German et al., 2016; Schmitt & Levine, 1977; Slaughter et al., 2006; Tomassetti et al., 2016). Once the regression weights of *what*, *when* and *how* are known, these numbers provide additional information to prioritise interventions that target the different components by considering the trade-off between their weights in the autonomy model and their potential effectiveness.

5.4 | Limitations and future work

Inspired by the policy-capturing method (Aarts et al., 1997; Aiman-Smith et al., 2002), our scenario-based task is ideal for maximising internal validity and experimental control. It may be criticised as too simplified or artificial. Participants in the studies were presented with scenarios that describe the allocation of decisions, but no actual decisions were made. While this is not necessarily a problem for testing the basic model, the scenario-based setup prevents researchers from answering some other interesting questions, for example, whether the match or mismatch between people's prior preferences and actual constrained decisions influences autonomy and agency experience. A related criticism is that in real life, unlike in our experiments, the determination of one component (e.g., the 'what') may further determine other components (e.g., the 'how'). While those cases do not necessarily weaken our model's ability to predict or explain (i.e., it would predict the effects of restricting both 'what' and 'how'), intervention designers should carefully consider whether restricting one component will implicitly restrict other components in certain situations (e.g., restricting travel mode may limit destination options).

For future research, we encourage researchers interested in our model to test it using other paradigms and methods, for example, experience sampling method, ecological decision-making tasks, or using conversational agents (e.g., chatbot) if the topic of AI and personal autonomy is of interest. By using more ecologically valid design, researchers have the opportunities to examine potential moderators of the core functional model, for example, the relationship between the decision-maker and the restricting agent, or the specific character of a human or AI agent. Another theoretically interesting moderator is whether the decision-maker perceives a restriction as legitimate or even internalises an external regulation on their behaviour (e.g., whether a person considers wearing face masks to be a rational and legitimate measure for combating COVID-19). According to the Self-Determination Theory, a highly internalised external regulation may invoke a feeling of volition or self-determination in complying with the regulation (Ryan & Deci, 2000). This may in turn result in preserved agentic experience.

Finally, our conclusions may be limited to the particular population from which our Prolific samples were drawn. While the age, gender and socioeconomic status of our sample were relatively diverse, most participants were white and from English-speaking countries. In these Western developed countries, there is a strong individualistic culture and personal autonomy is highly valued in society (Oshana, 2003). In more collectivistic cultures, such as in East Asian countries, there is a lot less emphasis on promoting personal autonomy and it is more acceptable to sacrifice autonomy for the sake of collective benefits (Helwig, 2006). It would be interesting to test whether the structure of our functional model holds for those populations and whether the negative effects of restricting all three components are less severe.

5.5 | Conclusion

To conclude, we observed that the effects of personal autonomy on human experiences of agency and goal motivation can be largely captured by restrictions in deciding what to do, how to act and when to act. The tested functional model provides crucial information about how experiences of agency and subsequently attitude and intention changes as a function of removing or adding one of the components from the autonomy equation. We hope and believe that our functional model of personal autonomy will stimulate future research to more precisely assess the effects of decision restrictions on people's experiences of being autonomous agents in today's fast-changing social, organisational and technological landscape.

ACKNOWLEDGMENTS

The reported research is part of the Alliance project "HUMAN-AI" funded by Utrecht University, Eindhoven University of Technology, Wageningen University and Research, and University Medical Center Utrecht. We thank Nil Akyüz for testing the experimental setup.

CONFLICT OF INTERESTS

The authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

Data, analysis scripts, pre-registrations and other materials can be found in the Open Science Framework (OSF) repository: https://osf.io/65xhf/

ETHICS STATEMENT

All participants in the studies signed informed consent forms and the experiments were approved by the Ethics Review Board of the Faculty of Social & Behavioural Sciences at Utrecht University.

ORCID

Chao Zhang https://orcid.org/0000-0001-9811-1881

REFERENCES

Aarts, H., & Elliot, A., (Eds.). (2012). *Goal-directed behavior*. Taylor & Francis. https://doi.org/10.4324/9780203869666

Aarts, H., Custers, R., & Wegner, D. M. (2005). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and Cognition*, *14*, 439–458. https://doi.org/10.1016/j.concog.2004.11.001

Aarts, H., Verplanken, B., & Van Knippenberg, A. (1997). Habit and information use in travel mode choices. *Acta Psychologica*, *96*, 1–14. https://doi.org/10.1016/S0001-6918(97)00008-5

Aiman-Smith, L., Scullen, S. E., & Barr, S. H. (2002). Conducting studies of decision making in organizational contexts: A tutorial for policy-capturing and other regression-based techniques. *Organizational Research Methods*, *5*, 388–414. https://doi.org/10.1177/109442802237117

Anderson, J., Rainie, L., & Luchsinger, A. (2018). *Artificial intelligence and the future of humans*. Pew Research Center. https://doi.org/10.1126/science.aat5991

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, *31*, 101–119. https://doi.org/10.1080/0305724022014322

Barlas, S. (2003). When choices give in to temptations: Explaining the disagreement among importance measures. *Organizational Behavior and Human Decision Processes*, *91*, 310–321. https://doi.org/10.1016/S0749-5978(02)00515-0

Bentham, J. (1789). *An introduction to the principles of morals*. Athlone.

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, doing, and controlling: Implicit sense of agency over somatosensory events. *Psychological Science*, *28*, 882–893. https://doi.org/10.1177/0956797617697693

Bostrom, N., & Sandberg, A. (2009). Cognitive enhancement: Methods, ethics, regulatory challenges. *Science and Engineering Ethics*, *15*, 311–341. https://doi.org/10.1007/s11948-009-9142-5

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist*, *14*, 319–325. https://doi.org/10.1177/1073858408317417

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.

Brookhouse, K. J., Guion, R. M., & Doherty, M. E. (1986). Social desirability response bias as one source of the discrepancy between subjective weights and regression weights. *Organizational Behavior and Human Decision Processes*, *37*, 316–328. https://doi.org/10.1016/0749-5978(86)90032-4

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality—social, clinical, and health psychology. *Psychological Bulletin*, *92*, 111–135. https://doi.org/10.1037/0033-2909.92.1.111

Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion changes the sense of agency in the human brain. *Current Biology*, *26*, 585–592. https://doi.org/10.1016/j.cub.2015.12.067

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, *56*, 809–825. https://doi.org/10.1177/0022243719851788

Corchado, J. M., Lees, B., & Rees, N. (1997). A multi-agent system "test bed" for evaluating autonomous agents. In *Proceedings of the first international conference on Autonomous agents*. (pp. 386–393). ACM. https://dl.acm.org/doi/pdf/10.1145/267658.267746

R Core Team (2020). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571. https://doi.org/10.1037/0003-066X.34.7.571

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95–106. https://doi.org/10.1037/h0037613

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674. https://doi.org/10.1126/science.2648573

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, *53*, 1024–1037. https://doi.org/10.1037/0022-3514.53.6.1024

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268. https://doi.org/10.1207/S15327965PLI110401

De Young, R. (1993). Changing behavior and making it stick: The conceptualization and management of conservation behavior. *Environment and Behavior*, *25*, 485–505. https://doi.org/10.1177/0013916593253003

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*, 114–126. https://doi.org/10.1037/xge0000033

Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge University Press.

Frankfurt, H. G. (1988). Freedom of the will and the concept of a person. In *What is a person?* (pp. 127–144). Humana Press. https://doi.org/10.1007/978-1-4612-3950-5_6

German, H., Fortin, M., & Read, D. (2016). Justice judgments: Individual self-insight and between-and within-person consistency. *Academy of Management Discoveries*, *2*, 33–50. https://doi.org/10.5465/amd.2014.0031

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*, 627–660. https://doi.org/10.5465/annals.2018.0057

Griffiths, P. E., & West, C. (2015). A balanced intervention ladder: Promoting autonomy through public health action. *Public Health*, *129*, 1092–1098. https://doi.org/10.1016/j.puhe.2015.08.007

Grüne-Yanoff, T., & Hertwig, R. (2016). Nudge versus boost: How coherent are policy and theory? *Minds and Machines*, *26*, 149–183. https://doi.org/10.1007/s11023-015-9367-9

Helwig, C. C. (2006). The development of personal autonomy throughout cultures. *Cognitive Development*, *21*, 458–473. https://doi.org/10.1016/j.cogdev.2006.06.009

Hoffstaedter, F., Grefkes, C., Zilles, K., & Eickhoff, S. B. (2013). The "what" and "when" of self-initiated movements. *Cerebral Cortex*, *23*, 520–530. https://doi.org/10.1093/cercor/bhr391

Ivanova, M., Bronowicka, J., Kocher, E., & Degner, A. (2018). Foodora and Deliveroo: The App as a Boss? Control and autonomy in app-based management-the case of food delivery riders (*No. 107*). Working Paper Forschungsförderung. http://hdl.handle.net/10419/216032

Jarrahi, M. H., Sutherland, W., Nelson, S. B., & Sawyer, S. (2020). Platformic management, boundary resources for gig work, and worker autonomy. *Computer Supported Cooperative Work (CSCW)*, *29*, 153–189. https://doi.org/10.1007/s10606-019-09368-7

Kamphorst, B., & Kalis, A. (2015). Why option generation matters for the design of autonomous e-coaching systems. *AI & SOCIETY*, *30*, 77–88. https://doi.org/10.1007/s00146-013-0532-5

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*, 366–410. https://doi.org/10.5465/annals.2018.0174

Korsgaard, C. M. (2009). *Self-constitution: agency, identity, and integrity*. Oxford University Press.

Krieghoff, V., Brass, M., Prinz, W., & Waszak, F. (2009). Dissociating what and when of intentional actions. *Frontiers in Human Neuroscience*, *3*, 3. https://doi.org/10.3389/neuro.09.003.2009

Kruglanski, A. W. (1996). Goals as knowledge structures. In P. M. Gollwitzer, & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 599–618). The Guilford Press.

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920951503

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*, https://doi.org/10.1177/2053951718756684

Lee, Z. W., Chan, T. K., Balaji, M. S., & Chong, A. Y. L. (2018). Why people participate in the sharing economy: An empirical investigation of uber. *Internet Research*, *28*(3), 829–850. https://doi.org/10.1108/IntR-01-2017-0037

Leotti, L. A., Iyengar, S. S., & Ochsner, K. N. (2010). Born to choose: The origins and value of the need for control. *Trends in Cognitive Sciences*, *14*, 457–463. https://doi.org/10.1016/j.tics.2010.08.001

Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy*, *1*, 35–42. https://doi.org/10.1353/bsp.2015.0000

Logg, J. M. (2017). Theory of machine: When do people rely on algorithms? *Harvard Business school working paper series #17-086*. http://nrs.harvard.edu/urn-3:HUL.InstRepos:31677474

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, *46*, 629–650. https://doi.org/10.1093/jcr/ucz013

Luck, M., & d'Inverno, M. (1995). A formal framework for agency and autonomy. In *Proceedings of the First International Conference on Multiagent Systems* (Vol. 95, pp. 254–260). AAAI.

Michie, S., & West, R. (2013). Behaviour change theory and evidence: A presentation to government. *Health Psychology Review*, *7*, 1–22. https://doi.org/10.1080/17437199.2011.649445

Mill, J. S. (1861). *Representative government*. Kessinger Publishing.

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

Nuffield Council on Bioethics (2007). *Public health: Ethical issues*. Nuffield Council on Bioethics.

Oshana, M. (2003). How much should we value autonomy? *Social Philosophy & Policy*, *20*, 99–126. https://doi.org/10.1017/S0265052503202041

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Raveendhran, R., & Fast, N. J. (2021). Humans judge, algorithms nudge: The psychology of behavior tracking acceptance. *Organizational Behavior and Human Decision Processes*, *164*, 11–26. https://doi.org/10.1016/j.obhdp.2021.01.001

Rosenberg, B. D., & Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, *4*, 281–300. https://doi.org/10.1037/mot0000091

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78. https://doi.org/10.1037//0003-066x.55.1.68

Ryan, R. M., & Deci, E. L. (2019). Brick by brick: The origins, development, and future of self-determination theory. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 6, pp. 111–156). Elsevier. https://doi.org/10.1016/bs.adms.2019.01.001

Sankaran, S., Zhang, C., Funk, M., Aarts, H., & Markopoulos, P. (2020). Do I have a say? Using conversational agents to re-imagine human-machine autonomy. In *Proceedings of the 2nd Conference on Conversational User Interfaces* (pp. 1–3). ACM. https://doi.org/10.1145/3405755.3406135

Schildt, H. (2017). Big data and organizational design–the brave new world of algorithmic management and computer augmented transparency. *Innovation*, *19*, 23–30. https://doi.org/10.1080/14479338.2016.1252043

Schmitt, N., & Levine, R. L. (1977). Statistical and subjective weights. Some problems and proposals. *Organizational Behavior and Human Performance*, *20*, 15–30. https://doi.org/10.1016/0030-5073(77)90041-1

Shapiro, A. (2018). Between autonomy and control: Strategies of arbitrage in the “on-demand” economy. *New Media & Society*, *20*, 2954–2971. https://doi.org/10.1177/1461444817738236

Slaughter, J. E., Richard, E. M., & Martin, J. H. (2006). Comparing the efficacy of policy-capturing weights and direct estimates for predicting job choice. *Organizational Research Methods*, *9*, 285–314. https://doi.org/10.1177/1094428105279936

Spruijt-Metz, D., Hekler, E., Saranummi, N., Intille, S., Korhonen, I., Nilsen, W., Rivera, D. E., Spring, B., Michie, S., Asch, D. A., Sanna, A., Traver Salcedo, V., Kukakfa, R., & Pavel, M. (2015). Building new computational models to support health behavior change and maintenance: New opportunities in behavioral research. *Translational Behavioral Medicine*, *5*, 335–346. https://doi.org/10.1007/s13142-015-0324-1

Tapal, A., Oren, E., Dar, R., & Eitam, B. (2017). The sense of agency scale: A measure of consciously perceived control over one’s mind, body, and the immediate environment. *Frontiers in Psychology*, *8*, 1552. https://doi.org/10.3389/fpsyg.2017.01552

Tauber, A. I. (2005). *Patient autonomy and the ethics of responsibility*. The MIT Press.

Tomassetti, A. J., Dalal, R. S., & Kaplan, S. A. (2016). Is policy capturing really more resistant than traditional self-report techniques to socially desirable responding? *Organizational Research Methods*, *19*, 255–285. https://doi.org/10.1177/1094428115627497

Väänänen, K., Sankaran, S., Lopez, M. G., & Zhang, C. (2021). Editorial: Respecting human autonomy through human-centered AI. *Frontiers in Artificial Intelligence*, *4*, 807566. https://doi.org/10.3389/frai.2021.807566

van Wissen, A. (2014). *Agent-based support for behavior change: models and applications in health and safety domains*. PhD thesis, VU University Amsterdam.

Vugts, A., van den Hoven, M., de Vet, E., & Verweij, M. (2020). How autonomy is understood in discussions on the ethics of nudging. *Behavioural Public Policy*, *4* 108–123. https://doi.org/10.1017/bpp.2018.5

Walker, R. L. (2008). Medical ethics needs a new view of autonomy. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, *33*, 594–608. https://doi.org/10.1093/jmp/jhn033

Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2012). The index of autonomous functioning: Development of a scale of human autonomy. *Journal of Research in Personality*, *46*, 397–413. https://doi.org/10.1016/j.jrp.2012.03.007

Zapparoli, L., Seghezzi, S., Scifo, P., Zerbi, A., Banfi, G., Tettamanti, M., & Paulesu, E. (2018). Dissecting the neurofunctional bases of intentional action. *Proceedings of the National Academy of Sciences*, *115*, 7440–7445. https://doi.org/10.1073/pnas.1718891115

Zhang, C. (2019). *Towards a psychological computing approach to digital lifestyle interventions*. PhD thesis, Eindhoven University of Technology.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhang, C., Sankaran, S., & Aarts, H. (2023). A functional analysis of personal autonomy: How restricting ‘what’, ‘when’ and ‘how’ affects experienced agency and goal motivation. *European Journal of Social Psychology*, *53*, 567–584. https://doi.org/10.1002/ejsp.2923