The Behavioral Mechanisms Behind Feedback – A Preliminary Model for Quantifying Cause-Effect Relationships

Carlo Stingl  
*University of Bamberg, carlo.stingl@uni-bamberg.de*

Sebastian A. Günther  
*University of Bamberg, sebastian.guenther@uni-bamberg.de*

Thorsten Staake  
*Otto-Friedrich Universität, thorsten.staake@uni-bamberg.de*

Follow this and additional works at: [https://aisel.aisnet.org/ecis2022_rip](https://aisel.aisnet.org/ecis2022_rip)

**Recommended Citation**  
[https://aisel.aisnet.org/ecis2022_rip/7](https://aisel.aisnet.org/ecis2022_rip/7)

This material is brought to you by the ECIS 2022 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2022 Research-in-Progress Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
THE BEHAVIORAL MECHANISMS BEHIND FEEDBACK – A PRELIMINARY MODEL FOR QUANTIFYING CAUSE-EFFECT RELATIONSHIPS

Research in Progress

Carlo Stingl, University of Bamberg, Germany, carlo.stingl@uni-bamberg.de
Sebastian A. Günther, University of Bamberg, Germany, sebastian.guenther@uni-bamberg.de
Thorsten Staake, University of Bamberg, Germany, thorsten.staake@uni-bamberg.de

Abstract

The ubiquity of information technology makes it possible to provide individuals with personalized, context-specific, and timely feedback at low marginal cost. This allows for highly scalable behavioral interventions that can support behavior change even for frequent, habitual, and incidental behaviors. Despite the well-documented potential of feedback interventions in improving personal and societal outcomes, the associated theoretical insights – in particular for real-time feedback on everyday behaviors – are sparse. To this end, this research-in-progress paper proposes a conceptual model to formalize cause-effect relationships between feedback interventions, the antecedents of behavior, and resulting behavioral changes by integrating important concepts from feedback intervention theory and other behavioral theories. Moreover, this paper outlines the planned validation of the model and research hypotheses with an empirical field study. Ultimately, the results are expected to inform the development of (future) information systems for behavior change by holistically explaining behavioral mechanisms associated with feedback interventions.

Keywords: Feedback Interventions, Theory Development, Structural Equation Modelling, Field Experiment.

1 Motivation

Everyday activities have collectively a large impact on societal outcomes, such as on people’s health (from eating to exercising, from washing hands to cleaning teeth) or our environment (from driving style to choice of transport, from shopping to showering). Such activities are often performed on “autopilot,” meaning automatically and without much deliberate cognitive processing. While this automaticity helps to reduce cognitive load and to manage the multitude of daily decisions more efficiently (Haith and Krakauer, 2018), it can also lead to adverse consequences for the individual (e.g., obesity resulting from habitual fast-food consumption) and society (e.g., inefficient use of scarce resources). With the emergence of information systems (IS) that collect and process sensor information, there is a wealth of data that can serve as the basis for digital interventions to promote conscious decision-making, e.g., by raising awareness of behavioral consequences or improving behavioral control. One of the most prominent behavioral interventions in IS research is the provision of feedback. Feedback provides recipients with tangible information on their behavior and thereby supports recipients to act in line with their goals. Indeed, a plethora of studies underline the potential of feedback interventions in overcoming behavioral issues for different domains (e.g., increasing physical activity (Normand, 2008), improving diets (Schembre et al., 2018), inducing resource conservation (Tiefenbeck et al., 2019), or reducing procrastination (Günther, 2021)).
Despite the growing use of feedback in IS research in this domain, there remain crucial gaps in the theoretical understanding of feedback interventions and their mechanisms, especially in the context of everyday activities. While behavioral models that try to explain the impact of feedback interventions can be classified into a) behavior-oriented and b) intervention-oriented models, they do not yet sufficiently inform the effective design of feedback-based interventions. For instance, behavior-oriented models such as the theory of planned behavior (Ajzen, 1985) focus on deliberate and reasoned behavior while neglecting habitual aspects, which limits their explanatory power for feedback interventions targeting routine activities. By contrast, intervention-oriented models such as feedback intervention theory (Kluger and DeNisi, 1996) or the behavior change wheel (Michie et al., 2011) describe cause-effect relationships on the impact of interventions, but do not weigh the role of specific behavioral mechanisms (e.g., behavioral control, correction of misbeliefs, attention, etc.) and their effects on behavior change. As a consequence, there is yet scarce knowledge on which characteristics are crucial for behavioral interventions to be most impactful. Given that our daily lives are now shaped by digital artefacts designed to influence our decisions, theory-based insights leading to more effective behavioral interventions could arguably have a great impact on addressing the grand challenges of our time (Burton-Jones et al., 2021).

Against this backdrop, we expand existing theories to untangle the mechanisms behind feedback interventions on routine activities. Our model is specifically designed for direct and immediate feedback IS can provide for everyday tasks that require little cognitive resources (e.g., driving, washing hands, exercising). As the effectiveness of feedback is task-specific (Kluger and DeNisi, 1996), it is important to note that our work does not address complex tasks in areas such as academic performance or the management of teams, where feedback is potentially ego-threatening (King, 2016) and for which an extensive body of research exists. We propose a structural equation model (SEM) to formalize influencing factors (antecedents) of feedback on behavioral performance to answer the following research question:

RQ1: How do antecedents influence behavioral performance in everyday tasks?

The influencing factors (antecedents) are inspired by feedback intervention theory and extended with certain mechanisms of influential behavior models we consider necessary for routine behavior. After the theory-driven formulation of the model, we describe how we plan to validate the model empirically. Experiments measuring real-world behavior might be especially valuable in this context, as self-reported behavior may be unreliable for routine activities (Sniehotta and Presseau, 2012; Gardner and Tang, 2014). For our validation, we aim to target an important routine everyday behavior where feedback interventions are likely to have a desirable impact on the health and life of recipients: washing hands. The activity is simple and behavioral aspects are accurately measurable by IS (e.g., length of hand wash, soap usage). We present an experimental design which not only validates the formulated SEM, but also shows the influence of feedback interventions on the structural relationships of the model. To achieve this, our research design aims at manipulating the outlined antecedents of behavior through the use of (feedback) interventions. After implementing the proposed research design in an upcoming field study (which remains work in progress and is not included in this paper), we then can answer the second research question:

RQ2: How are relationship of antecedents and behavior influenced by feedback interventions?

The paper is structured as follows. Section 2 sheds light on related research and crucial barriers for behavior change that feedback interventions may resolve, along with a more detailed description of feedback intervention theory which builds the heart of our model. Section 3 describes our approach, the developed model, testable hypotheses, and the planned validation of our model. Thereafter, the paper concludes by outlining expected contributions to theory and practice.

2 Related work

In the field of persuasive technology a growing amount of literature has been published on Behavior Change Support Systems (BCSS), with applications e.g. in health (Kelders et al., 2016), sustainability (Shevchuk and Oinas-Kukkonen, 2016) and for habitual activities (Karppinen et al., 2018). General
design principles and guiding behavior change theories for BCSS are suggested (Oinas-Kukkonen and Harjumaa, 2009; Oinas-Kukkonen, 2013). More specifically, the design principles in the Persuasive Systems Design Model (Oinas-Kukkonen and Harjumaa, 2009) present important aspects of feedback (e.g., self-monitoring, reminders and praise) as building blocks for IS design, but leave open a) how to combine them to an IS-delivered feedback intervention and b) how such feedback interventions are used most effectively to change behavior. We aim to add to this discussion by investigating and quantifying the behavioral mechanisms behind feedback. In selecting such behavioral mechanisms for our model, we draw on theories of behavior change mainly from behavioral psychology, which are also referred to as a guiding theories in the BCSS literature (e.g., Theory of planned behavior, Social cognitive theory, Goal setting theory (Oinas-Kukkonen, 2013)). Here, feedback interventions usually combine different types of manipulations to overcome crucial barriers for behavior change. In the following, we will highlight two important ways how feedback interventions typically overcome such barriers and subsequently concretize the research gap we will address.

First, feedback interventions typically aim at creating a sense of awareness and need for behavior change. This might be particularly relevant for everyday behaviors, as these are often characterized by a lack of deliberate decision-making (Kahneman, 2011). For instance, in the sustainability context, individuals are often not aware of the general consequences of their behavior, e.g., when consuming hot water in the shower or when deciding between two different products when grocery shopping. For most individuals, energy and CO₂ emissions are not salient – both are more or less invisible, their units of measure abstract, and costs are noticed, if at all, only after a delay (Faruqui et al., 2010). This “hidden nature” is also characteristic for many other everyday behaviors, such as in the health context (e.g., diet choices, physical activities) or financial context (e.g., need to invest in retirement plans). For many related decisions, people in addition often systematically discount future benefits of behavior change in favor of more immediate benefits (O’Donoghue and Rabin, 1999). Behavioral interventions may therefore mitigate undesirable consequences of this decision-making by simply creating awareness and pronouncing the relevance and need of behavior change (e.g., by providing knowledge about consequences and emphasizing social norms, see Goldstein et al., 2008). Indeed, many psychological models have integrated such antecedents to explain behavior, including Schwartz’s norm activation model (Schwartz, 1977), protection motivation theory (Rogers, 1975), or the theory of planned behavior (Ajzen, 1991).

Second, behavioral interventions often provide performance feedback to individuals, which might support individuals to behave in line with their aspirations. An important theoretical model behind the cause-effect relationships of feedback is provided by Kluger and DeNisi (1996), who analyzed over 100 papers to explore why feedback interventions sometimes lead to performance increase and sometimes to performance decrease. Feedback intervention theory postulates that individuals’ goals and standards are hierarchically oriented, and that individuals’ attention is usually directed at the medium level of this hierarchy (i.e., the task level). In response to feedback that receives attention, recipients regulate their behavior through comparisons of feedback to their corresponding goals and standard. If a discrepancy (gap) in these comparisons is detected, behavior is typically adapted to match the goals and standard by either increasing or lowering effort levels. If the discrepancy remains unresolved, attention may also shift to the lower level of the hierarchy (task detail level), which could trigger subsequent learning processes to improve performance, or to the upper level of the hierarchy (self-level), where individuals may question the relevance of the received feedback. The effectiveness of feedback also depends on task characteristics (Kluger and DeNisi, 1996), where feedback for tasks requiring low cognitive effort (e.g., the behaviors targeted by our model like washing hands or cleaning teeth) is hypothesized to have larger effects on performance than for cognitive demanding tasks (e.g., academic performance, see King, 2016).

Besides these mechanisms from feedback intervention theory, literature has also identified (i) that behavioral control and response efficacy are important antecedents in explaining behavior (see, e.g., Ajzen, 1991; Rogers, 1975) and (ii) that feedback may resolve individual perception errors on behavioral consequences (e.g., false belief of meeting one’s standard gaps, see Gerster et al., 2020). We argue that
feedback interventions may also positively influence these factors by providing individuals with tangible information about their past performance.

While there is a plethora of studies investigating the effects of feedback interventions on behavior (for an overview, see, e.g., Kluger and DeNisi, 1996; Loock et al., 2013), only a few studies specifically document how feedback interventions are influencing associated antecedents of behavior. A plausible reason for this is that popular intervention-oriented models, which provide meaningful insights over behavior-oriented models, are typically not mathematically formalized and therefore do not allow for such tests. From a practical and theoretical viewpoint, an advancement of such theoretical models, however, could be especially valuable. Specifically, feedback intervention theory postulates that context- and personality-specific variables might moderate the impact of feedback interventions (Kluger and DeNisi, 1996). As IS could automatically tailor feedback to recipients depending on such variables, theoretical insights might both advance our understanding and produce more effective behavior change systems. As a consequence, our goal is to contribute to research by developing a theoretical model (i) that integrates several important antecedents in the domain of feedback interventions and (ii) ultimately empowers theory-driven tests of so far only loosely formalized mechanisms.

3 Methodology

As a methodology for constructing our testable model, we follow the approach proposed by Ostrom (2005) based on the nested concepts of framework, theory, and model, which has been adopted in the IS literature (Burton-Jones and Grange, 2013; Baird and Maruping, 2021). It consists of the following steps: “(1) propose a high-level framework that defines the kinds of variables to be included in the theory and its associated metatheoretical assumptions, (2) apply the framework to a specific phenomenon, and (3) propose empirically testable models” (Burton-Jones and Grange, 2013, p. 640).

Following this three-step approach, we first select certain frameworks to serve as the underlying structure for theorizing on a general level. Such frameworks for our model building are control theory (Carver and Scheier, 1982) and social cognitive theory (Bandura, 1986, 1991). Control theory provides a conceptual framework for the adaptation of human behavior through self-regulation. Its basic construct is a “discrepancy-reducing feedback loop,” that compares a perceived state of behavior to a reference, resulting in adapted behavior in case a discrepancy between actual state and reference is detected. As a second framework, we use Bandura’s social cognitive theory (Bandura, 1986, 1991) as a comprehensive lens on cognitions and the social environment, which contains valuable constructs which we think influence the effectiveness of feedback, such as comparisons to social groups (norms), self-efficacy beliefs, and behavioral consequences. In the second step, the chosen high-level frameworks are applied to the specific phenomenon of feedback interventions. We view feedback intervention theory as such an application of control theory, as it contains similar mechanisms for behavior regulation in the form of iterative feedback-standard comparisons adapted to feedback interventions (Kluger and DeNisi, 1996). Feedback intervention theory further establishes the locus of attention as a key concept to explain the efficacy of feedback interventions. As the third step, we propose an empirically testable model that allows us to investigate a) how feedback interventions influence the antecedents of behavior and b) how subsequently the antecedents influence behavior.

3.1 Our proposed model

We propose a theoretical model for the influence of feedback on habitual everyday behavior, which we derive from related literature. Figure 1 shows our preliminary structural model, which is explained in the following. The accompanying measurement model is still in an early stage and out of the scope of this paper.
Perceived Standard Gap: Feedback intervention theory (Kluger and DeNisi, 1996) and preceding goal-oriented theories (goal setting theory, control theory) postulate that behavior is regulated by comparing behavioral performance (here: “Belief about Performance”) against a standard that originates from one’s ambitions and values (here: “Intended Performance”). If a discrepancy (gap) in this comparison is detected, behavior is typically regulated to match the standard by either increasing or lowering effort levels. We aim to formalize this mechanism in our model by including a perceived standard gap and connected constructs (Intended Performance, Attention, and Beliefs about Performance), which are explained in the following.

H1: The latent variable “Standard gap” has an impact on behavior.

Attention: A key assumption of feedback intervention theory is that „attention is limited and therefore only feedback-standard gaps that receive attention actively participate in behavior regulation” (Kluger and DeNisi, 1996, p. 259). To model this influence of attention on the relationship between standard gap and behavior, we include it as moderating variable. We also assume that the attention a behavior receives has a direct effect on it – even in the absence of feedback (see Schwartz et al., 2013).

H2: The latent variable “Attention” has an impact on behavior.

H3: The latent variable “Attention” moderates the relationship between “Standard Gap” and behavior.

Intended Performance: Intended performance refers to an internal standard that a person wants to achieve. Such standards play a role in goal-oriented theories such as goal setting theory (Locke and Latham, 1990). Intended performance is influenced by one’s norms regarding a behavior, where we distinguish between personal and social norms as both have been evaluated extensively before in the context of everyday behaviors (Allcott, 2011; Gaube et al., 2018; de Groot and Steg, 2008; Neighbors et al., 2004; Reid and Aiken, 2013; Roos and Hahn, 2019).

- Social Norm: A social norm is the perceived social pressure to perform, or not to perform, a specific behavior. Social psychology differentiates between two types of social norms (Schultz et al., 2007): Descriptive Norms describe the perception of how a social group actually behaves (Cialdini et al., 1991). Injunctive Norms describe the perception of what a social group thinks should be done, which manifests, e.g., in moral rules (Cialdini et al., 1991).
• **Personal Norm**: A personal norm is related to the individual moral obligation an individual has to perform, or not to perform, a behavior. The personal norm goes beyond social pressure as it depends on the value orientations an individual has with the target behavior. For instance, in the sustainability context, altruistic and biospheric value orientations seem to influence the personal norm (Roos and Hahn, 2019).

H4: The variable “Intended Performance” has an impact on the variable “Perceived standard gap.”

H5: The latent variable “Social Norm” has an impact on the variable “Intended Performance.”

H6: The latent variable “Personal Norm” has an impact on the variable “Intended Performance.”

**Belief about Performance**: Theories including self-monitoring and goal-directed behavior, such as control theory (Carver and Scheier, 1982), presuppose an evaluation of the current state of behavior, which is subsequently used for comparisons to a specific goal or standard. We include this internal evaluation of performance, as it is required to evaluate the standard gap in our model. In the case of handwashing, such beliefs could be that e.g., hands have been washed for long enough.

H7: The variable “Belief about Performance” has an impact on the variable “Perceived standard gap.”

**Beliefs on Consequences**: Many behavioral models include consequences of a person’s own behavior or action, e.g., as outcome expectancies in Bandura’s social cognitive theory (Bandura, 1986) or, specifically related to the welfare of others, as awareness of consequences in the norm activation model (Schwartz, 1977). We suggest that such consequences are also determinants of habitual behavior, especially as they very likely have an influence on the formation of habits (Frey and Rogers, 2014). In the case of handwashing, such beliefs on consequences might be assumptions that personal health is protected, or that frequent washing damages skin.

H8: The latent variable “Belief about consequences” has an impact on behavior.

**Perceived behavioral control**: A further important component is beliefs about the ability to perform a certain behavior correctly, e.g., wash hands for a certain amount of time. This construct is an important predictor of behavior in the theory of planned behavior (Ajzen, 1985) and related to Bandura’s concept of self-efficacy (Bandura, 1977), and thus we include it in our model.

H9: The latent variable “Perceived behavioral control” has an impact on behavior.

**Perception errors**: All constructs in the model so far measure perceived variables, which are subjective to the individual actor and may substantially differ from objective task performance. To account for possible deviations from perceived to objective behavior, we introduce the construct “Perception Errors” as the difference between perceived task performance (“Belief about Performance”) and objective (measured) performance. Such a perception error in the case of handwashing could be a person believing to wash hands for 20 seconds, whilst actually only washing for 10 seconds.

H10: The latent variable “Perception Errors” has an impact on behavior.

In summary, this model allows us to quantify cause-effect relationships between feedback interventions and their behavioral mechanisms.

### 3.2 Planned empirical validation of our model

To evaluate this structural equation model with behavioral interventions, we will conduct a $2 \times 2$ full factorial design with the conditions (no feedback, feedback) $\times$ (no norm, norm), as depicted in Figure 2. The first treatment condition is a social norm informing about the optimal duration of hand washing. The second treatment condition is an incrementing counter (in seconds) which starts automatically with the beginning of the handwashing procedure giving real-time feedback regarding the duration of handwashing. A display next to the sink, which is visible while washing hands, is responsible for
providing these interventions. Participants will be assigned randomly to these four conditions of approximately equal size. After washing hands (or having had the opportunity to wash hands), users are asked for their consent to participate in a follow up survey, measuring the constructs of our theoretical model. In case a user declines, their associated behavioral data will be discarded. Otherwise, survey data can be linked to previously measured real-world behavior (length of handwashing and usage of soap recorded by an unobtrusive IoT system, for details see Stingl et al., 2021). To minimize the effects of confounding influences such as the social desirability bias (Nederhof, 1985) on the behavioral data, we deliberately conduct the survey after the measurement of behavior.

Figure 2. Experimental design of the planned experiment

A control group (no feedback, no norm) makes it possible to estimate a reference model without treatment effects, and also to capture influences on the users’ behavior we cannot control for. The treatments groups extend this by potentially revealing resulting differences in the antecedents of behavior and, ultimately, causing differences in behavior. In the following, we describe the effects we anticipate from each intervention on behavior in the context of our research model.

The first intervention provides information about the recommended duration of washing hands. Thus, a participants’ internal social norm could be challenged, causing participants to either adapt or reject the new information. In case it is accepted, a new standard forms that influences the intended performance level. The behavior might subsequently be adapted by an evaluation of the updated standard gap. While studies testing similar interventions have reported small to no effect of poster-based interventions (Wearn et al., 2015; Lawson and Vaganay-Miller, 2019), we hypothesize this to be caused by missing attention directed towards standard gaps (a key mechanism also in the feedback intervention theory). To shed light on the underlying reasons, we therefore replicate prior research and hypothesize:

**H11:** Displaying the social norm will increase the influence of the latent variable “social norm” on “intended performance” and increase the influence of “standard gap” on behavior compared to the control group. It will result in only modest improvement, if any, in target behavior.

The second intervention which provides real-time feedback to participants displays an incrementing counter (in seconds) that starts automatically with the beginning of the hand washing procedure. This might direct the participants’ attention to the behavior itself and to the evaluation of the standard gap (without proposing a new standard). Beliefs about performance, for example the time already washed, are updated. Moreover, the intervention eliminates perception errors regarding length of hand washing and increases perceived behavior control. Thus, we hypothesize:
**H12:** Real-time feedback on performance will increase the influence of the latent variables “attention,” “standard gap,” “behavioral control,” and “perception errors” on the dependent variable compared to the control group. It will result in a considerable improvement in target behavior.

The third intervention provides both social norm and real-time feedback at once and potentially combines the mechanisms of the previously isolated interventions: a new standard for comparison for the standard gap is set via the displayed social norm, and the counter reduces perception errors, increases perceived behavioral control, and attention to the standard gap. We expect this condition to have the strongest effect on behavior and hypothesize:

**H13:** The combined intervention using a counter and norm will increase of the influence of the latent variable “Social Norm” on “Intended Performance” and increase the influence of the latent variables “attention,” “standard gap,” “behavioral control,” and “perception errors” on the dependent variable compared to the control group. It will result in a considerable improvement in target behavior.

Overall, with these interventions we aim to influence the antecedents of behavior in different ways and thus gain first insights on how our model can help to quantify behavioral mechanisms.

### 4 Conclusion and expected contribution

Feedback interventions delivered by IS have shown great potential in supporting individuals to improve both personal and societal outcomes. Despite their widespread use, however, little is known on the exact mechanisms by which feedback interventions induce behavior change, which currently hampers theorizing at a more general rather than situation-specific level. To shed light on these mechanisms, we have developed a preliminary model which incorporates many antecedents that we believe are crucial for explaining everyday routine behaviors. Specifically, our proposed structural equation model formalizes the influence of many behavioral antecedents under the lens of feedback intervention theory. Moreover, we have outlined research hypotheses and the planned validation of the model with an empirical field study we plan to conduct.

Our model is expected to contribute by quantifying the influence of different mechanisms of behavior change (e.g., attention towards behavior, self-regulation loops, and self-efficacy beliefs) and making them comparable against each other, highlighting which channels and mechanisms IS should target in general to facilitate behavior change most effectively. In addition, the model assesses how the specific feedback interventions of our study influence antecedents and ultimately behavior, demonstrating how an IS artefact can improve the outcomes of routine activities, which might also serve as blueprint for other behavior change systems. Limitations of our study include that the cross-sectional design only allows to measure short-term effects of feedback interventions and that the model is tailored specifically towards routine everyday behavior, which might fail to explain more complex tasks (e.g., academic performance).

Overall, this work will add to the yet scarce theoretical understanding of feedback interventions, as well as inform the design of effective feedback interventions in practice. With these insights, future IS could be more effective at helping individuals to change behavior and thus contribute to solving societal challenges.

### References


