Using feedback through digital technology to disrupt and change habitual behavior: A critical review of current literature

Sander Hermsen [a, *], Jeana Frost [b], Reint Jan Renes [a], Peter Kerkhof [b]

[a] Utrecht University of Applied Sciences, The Netherlands
[b] VU University Amsterdam, The Netherlands

Author Manuscript , please cite as:
Hermsen, S., Frost, J., Renes, R. J., & Kerkhof, P. (2016). Using feedback through digital technology to disrupt and change habitual behavior: A critical review of current literature. *Computers in Human Behavior, 57,* 61–74. doi:10.1016/j.chb.2015.12.023

Abstract

Habitual behavior is often hard to change because of a lack of self-monitoring skills. Digital technologies offer an unprecedented chance to facilitate self-monitoring by delivering feedback on undesired habitual behavior. This review analyzed the results of 72 studies in which feedback from digital technology attempted to disrupt and change undesired habits. A vast majority of these studies found that feedback through digital technology is an effective way to disrupt habits, regardless of target behavior or feedback technology used.

Unfortunately, methodological issues limit our confidence in the findings of all but 14 of the 50 studies with quantitative measurements in this review. Furthermore, only 4 studies tested for (and only 3 of those 4 found) sustained habit change, and it remains unclear how feedback from digital technology is moderated by receiver states and traits, as well as feedback characteristics such as feedback sign, comparison, tailoring, modality, frequency, timing and duration. We conclude with recommendations for new research directions.

Keywords

Digital technology; mobile and interactive technology; feedback; behavior change; habit change; habit disruption

Using Feedback from Digital Technology

to Disrupt and Change Habitual Behavior:

A Critical Review of Current Literature

1. Introduction

A variety of digital solutions to help us change detrimental or outdated habitual behavior have arrived on the market. These so-called *quantified self*-solutions, also known as *persuasive technologies*, aim to alter ingrained habits by presenting people with behavioral feedback through mobile and interactive devices and applications. These technologies can help individuals improve their health and the environment by increasing awareness and improving the self-regulation of behavior, something that does not come easily to us. Opportunities to incorporate such technologies in daily life have risen dramatically in recent years. In many nations, a great share of the general populace owns a smartphone or other kind of smart device and seems willing to use technology to change unwanted behaviors. For instance, more than 69% of US citizens track at least one health behavior, with 14% using a specialized tracker (Fox & Duggan, 2012). Manufacturers are jumping on this bandwagon, offering new ways to measure behavior, e.g. through Apple's Research Kit (Moynihan, 2015).

Few of these quantified self-products have been tested in controlled circumstances (Cowan, Bowers, Beale, & Pinder, 2013). Moreover, most solutions lack scientific evidence, with positive anecdotal reports from practice comprising the basis of our understanding (Cowan et al., 2013; Schoffman, Turner-McGrievy, Jones, & Wilcox, 2013). As yet, the potential of digital technology to disrupt and possibly even change habits through feedback on habitual behaviors remains unclear.

This paper addresses this gap in the literature by presenting a review of existing studies on the use of feedback generated by digital technology to disrupt and change automatic, habitual behaviors. This review adds to the current debate by providing an overview of existing evidence, accentuating and addressing gaps in current knowledge and laying an evidentiary foundation for digital technology solutions aimed at habit change.

To do so, we first assess the drawbacks of habitual behavior and the strategies that may be applied to disrupt undesired habits. Second, we then discuss the role of self-monitoring in habit disruption and the role feedback from external sources can play in self-monitoring. In the third section, we look at known influences of feedback efficacy, and consider whether insights into the effect of feedback on habitual behavior in general are valid when applied to feedback delivered through digital technology. Finally, we review findings on the use of digital technology that utilizes feedback and suggest avenues for future research.

1.1 Habitual behavior

In everyday life, habits, commonly defined as "behavior (...) prompted automatically by situational cues, as a result of learned cue-behavior associations" (Wood & Neal, 2009, pp. 580; Gardner, 2014, p.1), help us to come to terms with the enormous complexity of everyday life. However, some of the biggest threats to personal and planetary wellbeing are direct consequences of our habitual behavior. The cue-response-chain of a strong habit is a rigid structure, which overrides contradictory behavioral intentions (Verplanken & Faes, 1999; Verplanken & Wood, 2006). This may lead to undesired results when cue-response-pairs have a satisfying short-term effect but lead to damaging consequences in the long run, as with snacking or alcohol abuse. Furthermore, since habits do not take into account current context, changed circumstances may render habits unproductive for contemporary life, even though the behavior may have led to rewards in the past.

Because habitual behavior circumvents active consideration of the current context, it is hard to change habits using interventions aimed at controlled processing, e.g. through persuasive messages (Verplanken & Wood, 2006; Jager 2003). One powerful strategy to disrupt habits is therefore to change the circumstances so that habit cueing does not occur (Verplanken & Wood, 2006) or to alter the external cues that lead to habit execution (e.g. in Aarts & Dijksterhuis, 2003). However, these strategies have practical difficulties, since manipulating or avoiding cues is often impossible (Quinn, Pascoe, Wood, & Neal, 2010) and not always seen as ethical, because receivers may not always consciously notice the manipulations, which places their consequences outside the reach of conscious scrutiny (Verbeek, 2006).

1.2 Disrupting and changing habitual behavior by self-monitoring and feedback

The automaticity of habitual behavior means that execution is often at least partially unconscious and may start without conscious intent (Bargh, 1994). Therefore, one way to disrupt undesired habits is to bring habitual behavior and its context to (conscious) awareness. Self-monitoring, the procedure by which individuals record the occurrences of their own target behaviors (Nelson & Hayes, 1981), enables perception of our own behavior and adaption to the current context. Thus, self-monitoring leads to decreases in unwanted behavior (Quinn et al., 2010).

Unfortunately, self-monitoring is difficult for even the most motivated individual (Wilson, 2002). For example, there is often a discrepancy between self-reported and actual performance, as shown in diverse behaviors such as calorie intake (Lichtman et al., 1992), weight and BMI - especially in overweight participants (Pursey, Burrows, Stanwell, and Collins, 2014), the amount of exercise (Lichtman et al., 1992), actual versus perceived water use (Hamilton, 1985; Millock & Nauges, 2010), and even the reporting of relatively stable personal data such as height (Pursey et al., 2014).

Accurate self-monitoring is greatly improved by personalized information from external sources (Kim et al., 2013; Li, Dey, & Forlizzi, 2010). The intentional delivery of such information about performance or behavior (or about the impact of one's performance or behavior) in order to facilitate behavior change is commonly referred to as *feedback* (Van Velsor, Leslie, & Fleenor, 1997, p. 36). In this review, we adopt the definition of feedback offered by Kluger and Denisi (1996), seeing feedback as "actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance"[1].

The beneficial effect of feedback on performance has been established in a range of fields. In education, the role of feedback is especially well established. Hattie and Timperley (2007) performed a synthesis of meta-analyses of feedback in educational contexts and reported an average effect size of 0.79 for feedback interventions, almost twice the average effect size of general educational interventions (0.40). This implies that feedback interventions in general are not only capable of disrupting undesirable habits, but can also play a significant role in changing those behaviors. Similarly, feedback has been shown to be effective in an increasing range of

[1] This definition excludes non-task-related feedback ("he just does not like you"), and intrinsic, task-generated feedback (e.g. getting coffee from a machine and seeing that your coffee cup is full), whilst including feedback on *how* a task is performed (e.g. "you kicked the ball with the tips of your toes; you should have used the instep" in football training).

controlled studies regarding both health (Gardner et al., 2010) and sustainability (Darby, 2006; Froehlich, Findlater, & Landay, 2010; Fischer, 2008).

1.3 Feedback on behavior through digital technology

Direct, instant feedback used to be difficult to deliver regularly on a large scale. The delivery of feedback was restricted to either distant, impersonal media such as utility bills and letters, or cost-intensive face-to-face communication with trained personnel. The advent of mobile and interactive media has changed that. In recent years, technological developments have enabled a surge of behavior-changing interventions. A range of mobile apps, wearable devices, web-based platforms and in-home displays give us feedback on our behavior and monitor behavior that previously remained hidden. There are apps and wristbands to support us in physical exercise, applications for weight loss, in-home displays to encourage us to use less energy, etcetera.

Already, more than half of smartphone users gather health-related data with their phone, one in five has installed at least one health-behavior related app (Fox & Duggan, 2012) and one in ten Americans owns some sort of automatic activity tracker (Ledger & McCaffrey, 2014). Similarly, many European countries aim to achieve smart energy meter installation in every home by 2020 (Faruqui, Harris, & Hledik, 2010).

Digital technology can offer constant, real-time updates on our progress, powered by sensitive measuring devices, often worn on the body. The widespread use of sensing systems means that automatically generated data about the undesired behaviors can be made available, without the need for possibly problematic self-reporting. Monitoring devices can be used for a range of data-gathering causes including health statistics like heart rate, blood pressure, and blood sugar (Verplanken & Wood, 2006) and environmentally important data on energy use (Verplanken & Wood, 2006; Froehlich, Findlater, & Landay, 2010).

Besides data generation, digital technology can offer habit-disrupting cues such as light signals, buzzes, beeps, and push messages. Digital technology is not only useful to present users with evaluations of past behavior ("reflection-on-action"); because of the ubiquity of mobile and handheld devices, digital technology offers an unprecedented opportunity for "reflection-in-action" (Schön, 1984), the analysis of behavior as it occurs.

The availability of interactive displays provides ample opportunity for new types of feedback. A power socket may be enhanced to report energy use (Heller & Borchers, 2011), a shower head can give us feedback on water use or shower time (Andler, Woolf, & Wilson, 2013), or a power cable can move around as if in agony if connected devices are left in stand-by mode (Laschke, Hassenzahl, & Dieffenbach, 2011).

Digital technology has a number of distinct advantages over human persuaders. Devices can be (irritatingly) persistent, guarantee greater anonymity and have access to areas where people are not welcome (e.g. the bedroom or bathroom) or unable to go (e.g. inside clothing or household appliances). Moreover, digital technology is relatively easy to replicate, distribute and tailor to specific needs (Fogg, 2003). However, there are some disadvantages: digital technology is a lot easier to ignore or shut down than messages delivered by human persuaders. Furthermore, digital technology solutions are easily forgotten, lost or otherwise misplaced. For example, over half of those that have owned a wearable fitness tracker no longer use it, and a third of the users quits use in the first six months after purchase (Ledger & McCaffrey, 2014). Yet, in providing automatically delivered feedback for habit change, the benefits of digital technology may very well outweigh the disadvantages.

1.4 How feedback works: Mechanisms underlying feedback efficacy

Control theory provides insight into the mechanisms underlying the effect of feedback (Carver & Scheier, 1985). According to control theory, reflective behavior change processes are reminiscent of a thermostat. When looking to change their behavior, people compare their performance to a behavioral goal. When a discrepancy is noted, given enough motivation, opportunity, and the right abilities, people will attempt to reduce this discrepancy. The efficacy of this regulatory cycle is moderated by three executive function skills (cf. Hoffman, Schmeichel, & Baddeley, 2012): keeping a goal salient in working memory or bringing the goal back to working memory when needed; the ability to inhibit undesired automatic responses; and the ability to switch between tasks or mental sets.

Feedback supports reflection by increasing knowledge and awareness of behaviors and their impacts. Many behaviors are of such automaticity, that their performance is at least partly subconscious. Knowing *that* and *when* a habit occurs opens up possibilities for behavior change. Feedback also enables us to compare the consequences of our behavior to our current goals and adapt when the behavior does not fit the context. Furthermore, it also serves to increase general self-awareness, which in turn increases our capabilities to inhibit undesired behaviors (Alberts, Martijn & De Vries, 2011).

Feedback also has motivational consequences. We are driven by motivations to approach experiences that are expected to be pleasurable, and avoid unpleasant experiences (Elliot & Covington, 2001; Higgins, 1997). Both the negative emotions caused by an observed increasing discrepancy between goals and performance, and the positive emotions caused by a decreasing discrepancy, can increase our motivation to reach our goals (Carver & Scheier, 2011; Deci, Koestner & Ryan, 1999). Furthermore, among competing behaviors, those supported by feedback are given priority over those without feedback (Northcraft, Schmidt, & Ashford, 2011).

1.5 Factors moderating feedback efficacy

In a meta-analysis of 607 studies, Kluger and DeNisi (1996) found that, generally speaking, two thirds of all feedback interventions increased performance. However, the remaining third of the interventions had an opposite, detrimental effect on performance. Importantly, this means that even though we can expect a habit-disrupting effect from well-designed feedback interventions, this does not automatically signify that the feedback intervention will lead to change in the desired direction.

Furthermore this suggests that an interplay of receiver states and traits on the one hand, and feedback properties such as content (e.g. sign, comparison and level of detail), timing, modality, frequency, duration, and presentation on the other, influence feedback effectiveness (Fischer, 2008). The moderating effects of both receiver traits and states and feedback properties will be discussed below.

1.5.1 Interpersonal and intra-personal differences

Feedback efficacy is moderated by all kinds of characteristics of the feedback receiver, in an interplay of stable and more dynamic factors. A great deal of the expected moderators is stable and relatively uncontrollable, such as socio-economic status (e.g., Maitland, Chambers & Siek, 2009: affluent participants seem to benefit more from feedback interventions than poorer participants) and gender (e.g. Guadagno & Cialdini, 2007; Ho et al., 2013).

In any self-control mechanism, executive control capabilities play an important role, such as the capacity for self-regulation. Differences in personality and context determine the degree to which an individual is capable of exercising such control (Baumeister & Heatherton, 1996;

Braverman, 2008; Kuhl, 1985). In addition, self-regulating capacity is in finite supply (Baumeister et al., 1998).

Feedback efficacy is also influenced by relatively fleeting states such as high initial engagement with the target goal, strong motivation or a high perceived self-efficacy (Bandura, 1997). Self-regulation processes are cyclical in nature (Bandura, 1997; Zimmerman, 1998). This indicates that high initial motivation leads to a greater feedback effect, which in turn leads to increased motivation (e.g., Geister, Konradt, & Hertel, 2006). Similar cyclical effects can be found for self-regulatory skills and perceived self-efficacy (e.g. Donovan & Hafsteinsson, 2006; Multon, Brown & Lent, 1991).

To date, there is little or no evidence on whether these intra- and interpersonal factors that are generally known to influence feedback efficacy, such as motivation and perceived self-efficacy towards the goal, self-regulatory capabilities, and demographic and socio-economic factors, have different effects on the efficacy of feedback when it is delivered through digital technology. Since the latter is generally delivered in an individual context and not within the social setting of interpersonal feedback, the effect of feedback through digital technology might rely on capabilities and motivation of the receiver more than with interpersonal feedback.

1.5.2 Feedback properties

Paying attention to carefully crafting the timing, delivery, and content of the feedback can enhance the effectiveness of feedback interventions. In an extensive review of feedback on household energy use, Fischer (2008) indicates that high frequency feedback delivered over a long period by computerized and interactive tools provides an advantage in feedback effectiveness. There are a number of feedback properties that may affect effectiveness, including *technology, content, timing, modality, duration, frequency,* and *presentation and user* *experience.* Generally, the largest effects can be expected from detailed, positively framed, concurrent feedback ('reflection-in-action'), delivered continuously or on-demand through more than one modality, during a long period.

Technology. Feedback can be delivered through many different technological channels, ranging from websites and smartphone apps to wearables and in-home displays. The possibility to deliver well-designed and automatically tailored, in-action, frequently delivered feedback over a long period of time is one of the perceived strengths of digital, interactive technology. Because behavior often is measured directly, a direct response is possible, and the all-pervasive use of smartphones and other technologies means instant delivery on a large scale is relatively easy.

Each form of the technology has its advantages and disadvantages as a source of feedback. For example SMS text messages, a well-researched and generally considered effective means of feedback delivery (Hall, Cole-Lewis, & Bernhardt, 2015), are difficult to deliver at the very moment the behavior occurs because of time lag. This delay can severely disrupt performance, which may in some cases have negative consequences on behavioral fluency (Bittner & Zondervan, 2015). Furthermore, text messages can only deliver content of limited length (usually about 160 characters). On the other end of the spectrum, wearable activity trackers can do real time tracking of behavioral data, and are capable of on-demand or continuous delivery over a range of sensory channels without limits to the richness of the data (Yang & Hsu, 2010).

Content. Tailoring content to fit receiver characteristics can be expected to affect feedback effectiveness. Ample evidence from the literature shows that tailoring message content to meet recipient motivation, traits, abilities and preferences increases the effectiveness of such messages (e.g. Noar, Benac, & Harris, 2007; Noar, Harrington, Van Stee, & Aldrich, 2011; Ivers et al., 2012; Kaptein, De Ruyter, Markopoulos, & Aarts, 2012). Such tailoring may encompass utilizing negative, positive or neutral feedback (i.e. feedback sign); offering social, historical or normative comparisons (or no comparison at all); and increasing or decreasing level of detail.

Timing. There has been substantial research on the effect of feedback timing on learning (Hattie & Timperley, 2007, p. 98). Specifically, reflection-in-action can be expected to be more effective than reflection-on-action. For instance, in electricity use, direct, short delay feedback on energy usage generally leads to a 5–15% reduction in consumption, and indirect, long delay feedback leads to a reduction of 0–10% (Darby, 2006).

Modality. Selecting optimal delivery through visual, auditive, or tactile channels, or a combination of channels, increases feedback effectiveness (Hoggan, Crossan, Brewster, & Kaaresoja, 2009; Warnock, McGee-Lennon, & Brewster, 2011; Braverman, 2008). An optimal modality choice depends on the possibility of disruption and the need for detail. The visual mode is more disruptive than the auditory, which is in turn more disruptive than tactile feedback. Similarly, visual feedback can contain more detailed information than auditory, which in turn has more capacity for detail than tactile feedback.

Frequency and duration. Frequency and duration of the feedback intervention also influence feedback effectiveness. In general, the more frequent the feedback is delivered, over a longer period of time, the more the intervention will contribute to behavior change. The benefits of more frequent feedback are limited by cognitive capacity: as long as the frequency of the feedback does not overwhelm an individual's cognitive resources, more feedback is better (Lam et al., 2011). Current technological developments, especially those that concern use of mobile and interactive platforms, make it possible to circumvent these limitations and easily deliver much more frequent or even continuous feedback, with infinite durations. In theory, this should increase feedback effectiveness.

Presentation and user experience. Research in web design (Tuch et al., 2012), typography (Larson & Picard, 2005) and usability (Tractinsky, Katz, & Ikar, 2000) suggests that visual design aspects and aesthetics determine the attitude towards a design as well as the perceived ease of use (but not actual use). Consequently, users will feel more beneficial towards an aesthetically pleasing intervention and will be more inclined to persevere in using it. Moreover, a clear design might aid in emphasizing important information, personalizing the feedback and improving the fluency of feedback. However, the design and presentation of the feedback and technology must also fit participants' goals. For example, research on the design of glucometers suggests that the desired look and feel depends on context; users favor a more “medical” appearance when passing through customs on transatlantic flights and inconspicuous or sporty looks in day to day life (O'Kane, Rogers, & Blandford, 2015).

1.6 Reviewing the effects of feedback delivered by digital technology

Feedback through digital, interactive technology can have two beneficial effects on habitual behavior. Firstly, it can disrupt the automatic execution of the habitual behavior, making it available for conscious scrutiny. Secondly, feedback can lead to durable behavior change. Given the extensive evidence for the beneficial effect of feedback on habitual behavior change in general (e.g. Brug et al., 1998; Fischer, 2008; Hattie & Timperley, 2007; Ivers et al, 2012, Kluger & DeNisi, 1996), and the aforementioned benefits of digital technology over more traditional forms of feedback delivery, one assumption in this work is that feedback delivered by digital technology is at least as effective as 'regular' feedback in disrupting undesired habits.

Furthermore, based on literature on feedback on habitual behavior in general, feedback delivered through a well-chosen digital technology appears well suited to increase the chances of durable, lasting behavior change.

However, the fact that feedback through digital technology is delivered without the intervention of a human source might influence its effect, e.g. because of the lack of social pressure. Similarly, the effects of receiver moderators such as motivation and perceived self-efficacy are likely, but not certain, to be similar to those reported for feedback in general (the more motivation or the higher the perceived self-efficacy, the more effect of feedback can be expected).

The current review provides an overview of recent original studies that look into the effect of feedback through digital technology on undesired habitual behaviors. This review provides an analysis of the efficacy of such feedback to both disrupt and durably change habitual behavior. Furthermore, the review evaluates the effects of interpersonal and intra-personal differences; technology choice; and feedback properties: technology, content, timing, modality, duration, frequency, and presentation and user experience, on feedback efficacy.

2. METHOD

A combined search of the databases PubMed, PsychInfo, EMBASE and Web of Science was performed with the following set of search terms: (habit* OR habitual behavior) AND (persuasion OR behavior change OR habit disruption) AND (feedback OR self-monitoring) AND (persuasive design OR persuasive technology OR digital technology). This search resulted in 993 results. The ACM Digital Library and the IEEE Xplore Digital Library were searched, using the search terms "feedback AND persuasive AND habit". This search yielded 416 results from ACM/DL and 233 results form IEEE/Xplore; these results included peer-reviewed journal papers as well as conference proceedings.

Abstracts from both result sets were checked for relevance. From these, 101 publications with relevant and ambiguous abstracts were retained. Papers cited in included articles were checked for eligibility. Ancestry searches were performed on the included articles through Google Scholar, to retrieve more recent articles building upon the original work. From these searches, a further 35 primary publications were included. This resulted in a set of 136 primary sources.

From this set, 69 original papers matched the following inclusion criteria:

- The research has the primary purpose of changing habitual behavior, either increasing or decreasing the behavior or stopping the behavior altogether. Habit is operationalized as recurring behaviors with some degree of automaticity (Wood & Neal, 2009)
- Digital technology has to be used as the primary means of achieving behavior change
- The digital technology must use a tailored feedback mechanism delivered by (an) external agent(s) to provide information regarding task performance
- The research must encompass some form of analysis of the effect of the intervention on the targeted behavior, be it qualitative or quantitative.
- Because of rapid developments in the field of digital technology, only papers from the last decade (2004 and later) were included.

All analyzed papers are included in the reference list and marked with an asterisk (*). One included paper reported three relevant studies (Nakajima & Lehdonvirta, 2013) and two papers reported two relevant studies (Connelly et al., 2006, and Stienstra, Wensveen, & Kuenen, 2011), all of which were separately scored. This resulted in a final set of 72 studies.

The broad range of dependent variables, feedback intervention technologies, and research methods applied in the included papers made it impossible to conduct a meta-analysis of results in such a way that it would produce reliable and valid insights (Borenstein, Hedges, Higgings, & Rothstein, 2009; Quintana, 2015). Consequently, a systematic review with a descriptive analysis (Garg, Hackam, & Tonelli, 2008) of the literature was performed. Even though, when compared to a meta-analysis, a systematic literature review has more limited possibilities to derive general conclusions, this approach is able to shed light on the general direction of effects, as well as identify gaps in the literature (ibidem). Furthermore, conducting a systematic literature review enables us to incorporate results from qualitative studies, which would not be possible in a meta-analysis.

We thematically classified target behaviors of the intervention, feedback technology, feedback characteristics (content (feedback sign, comparison, and level of tailoring), timing, modality, frequency, duration, data source), and the availability of visual examples of the design and provided feedback. For each intervention, number of participants, independent and variables, analysis method, results, and possible methodological concerns were scored.

The included studies covered a range of dependent variables, varying from energy consumption to motor skills and physical activity. A list of the occurrence of each category of dependent variable is included in table 1. A full list of included studies, including target behaviors, feedback content, characteristics, dependent and independent variables and measurement methods is available as an online supplement.

Table 1: dependent variables

24	energy and water consumption
11	motor skills (speed skating, posture, violin playing, tooth brushing)
10	healthy eating and weight loss

9	physical activity
6	driving
3	general wellbeing
3	waste reduction
2	break taking and resuming work
9	other (social feedback, bookshelf ordering, IQ training, printing behavior, medication adherence, overfilling water cookers, transport mode choice)

3. RESULTS AND DISCUSSION

In this section, we first discuss the consequences of the diverse methodological approaches, followed by an analysis of review results ordered by theme – general effects of feedback on disrupting and changing habitual behavior, the effect of receiver characteristics, and the effects of different feedback technologies and characteristics. Finally, we discuss a few insights that transpired from qualitative results that were not based on a pre-posed hypothesis.

3.1 Methodological issues

The broadness of the range of studies included in this review is reflected in the different methodological approaches used. Of the 72 studies included in this review, three studies took place under controlled (laboratory) circumstances, 20 were field studies (7 of which were set up as a randomized controlled trial), and 49 studies tested a prototype or design. With regard to methods of analysis, 21 studies used qualitative analysis, mostly user experience studies describing interactions with designed prototypes. 50 studies utilized some form of quantified measurement and analysis, in 15 cases together with qualitative measures. In one paper, data gathering and analysis were described so poorly, that it remained unclear which research methodology was used.

Each form of research design and method of analysis has its own unique merits to the generation of knowledge. However, in every research design, reliability and validity should be well thought-through, to prevent experimental artifacts such as the Hawthorne effect – mere observation enhancing performance (cf. McCarney et al., 2007) –, demand characteristics – participants' interpretation of what is expected of them (Orne, 1962), or unforeseen events influencing performance – such as seasonal influences on energy use that may eclipse the effect of a feedback intervention. In general, quantitative studies that include (active) control groups, pre- and post-test measures, and use a fitting statistical test with ample power (Maxwell & Delaney, 2004, p. 56–59) are better suited for this. In qualitative study designs, a well-structured data collection and analysis strategy is necessary to reduce the chance of cherry-picking precisely those results that fit the hypothesis (Patton, 1990).

Most of the included quantitative studies did not meet these criteria. 33 of 50 quantitative studies did not report a strategy of dealing with experimental artifacts such as demand characteristics or unforeseen external moderators. Of the 50 quantitative studies, 30 studies were analyzed using statistical testing, yet only 8 out of these 30 studies showed sufficient statistical power for the sort of analysis performed. This is important, since low statistical power implies a large chance of type I and II errors (Cohen, 1992). Furthermore, low statistical power combined with a significant result dramatically increases the chance of an overestimation of intervention effects (Gelman & Carlin, 2014). In total, only 14 out of 50 studies with some sort of quantitative measurements had sufficient statistical power plus an experimental design that would prevent the occurrence of the most common experimental artifacts.

The 21 qualitative studies included in this review were all of sufficient rigor to avoid cherry picking in results. Most studies used a form of structured interviewing as data collection method, and reported some sort of systematic appraisal of the results. No qualitative studies had obvious methodological shortcomings.

We focus our analysis on those studies that meet all criteria mentioned above, both utilizing qualitative and quantitative methods. Subsequently, we will mention descriptive results from studies that did not meet all these criteria, with a corresponding caveat.

3.2 The effect of feedback through digital technology on *disrupting* habitual behavior

The effect of feedback through digital technology on disrupting habitual behavior is generally confirmed by our analysis. Of the 72 studies included in this analysis, 59 studies show a beneficial effect of feedback on disrupting habitual behavior. 13 of 14 studies with well-set up quantitative experimental designs and ample statistical power report significant results. 25 studies show a beneficial effect based on qualitative measurements, including observation reports, interviews and other user experience measures. Furthermore, from the remaining 37 quantitative studies, 32 studies report descriptive data that point in the direction of hypothesis. Of all studies that report a beneficial effect, five studies found this effect to be partial, i.e. not occurring in every expected condition.

Thirteen of fourteen experimental studies prove the beneficial effect of feedback through digital technology on a broad range of habitual behaviors. Feedback increased fruit consumption (Bech-Larsen & Grønhøj, 2013), safer driving behavior (Donmez, Boyle, & Lee, 2008; Maltz & Shinar, 2004), motor learning (Lieberman & Breazeal, 2007) and posture training (Epstein et al., 2012), lowering eating rate (Ford et al., 2010), increasing physical activity (Hurling et al., 2008; Schulz et al., 2014), weight loss (Pellegrini et al., 2012; Schulz et al., 2014), limiting computer use (Van Dantzig, Geleijnse, & Van Halteren, 2013), shower use (Willis et al., 2010), and electricity consumption (Jain, Taylor, & Peschiera, 2012; Wood & Newborough, 2003; Vassileva, Odlare, Wallin, & Dahlquist, 2012).

One well-designed quantitative study reported a null effect. The lack of effect in this study, in which participants could volunteer to join a home energy reduction intervention (Alahmad et al, 2012), could be ascribed to a ceiling effect caused by participant self-selection, such that only highly motivated participants that already performed many energy-saving behaviors took part. This could prove a limitation of the efficacy of feedback interventions: when participants are already performing the behavior in some way, there is a limit to habit change coming from feedback.

Seven qualitative studies reported no effects or even a contrary effect of feedback on behavior change. One study on waste disposal (Comber & Thieme, 2013) and a study on electricity usage (Hargreaves et al., 2010) found that although no behavior change was registered, knowledge about which behaviors were desirable and which less so did increase. In two studies, participants did not understand the manipulation (Gyllensward, Gustafsson & Bång, 2012; Kim et al., 2008). One further study (Nakajima & Lehdonvirta, 2013) on utilizing feedback to encourage a certain ordering of books on a bookshelf, led participants to play around with the installation, with inverse effects. Inverse effects were also found in a study on taking breaks at work, where participants used social activity feedback to avoid colleagues or to find empty rooms for meetings (Kirkham et al., 2013). This, too, may be a limitation of feedback: receivers may not perceive the feedback as a cue towards the target behavior. Studies by Katzeff et al. (2012) on energy use in the office, and Strengers (2011) on energy and water consumption show how feedback may not per se lead to behavior change, but may in the latter case also cause post-hoc rationalizations of the undesired behavior.

Finally, four quantitative studies found null results; however, all four studies (Cowan et al., 2013; Rodgers & Bartram, 2011; Pereira et al., 2012; Quintal, Pereira & Nunes, 2012) suffered from a lack of statistical power, so their null finding may very well be due to small sample sizes, since descriptive results in all studies do show a small positive effect of the reported interventions.

Where possible, we calculated effect sizes of quantified measurement methods for comparison (table 2). 28 studies either reported effect sizes or presented their data in such a way that effect sizes could be calculated. Even though the broad range of dependent and independent variables used in the reviewed studies make direct comparison in the form of a meta-analysis unfeasible, an overview of effect sizes listed could in theory serve as an indication of effect sizes to be expected in feedback interventions on habitual behavior.

Because of the methodological issues in the greater part of these studies, the reported effect sizes should be used with extreme caution. Low statistical power, especially, increases the chance of inflated effect sizes (Gelman & Carlin, 2014), which would give at least a partial explanation of the size of the effects found in many studies in this review.

Table 2: Effect sizes (reported or calculated)

Study	*Effect Size (Cohen's d)*	*Dependent variable*	*Participants*	*Analysis*[1]	*Issues*[2]	*Field*[3]

Hurling et al., 2008	3.022	physical exercise	70	Other		b
Hoggan et al., 2010	2.5201	IQ training	9	a	a	a
Stamopoulos et al., 2014	2.3528	buying domestic products	32	b	a	a
Chang et al., 2008	2.129	Brushing teeth in children	13	c	a, b	a
Spelmezan, 2012	1.9604	snowboarding skill	10	a	a, b, c	a
Bruns Alonso et al., 2014	1.6101	toothbrushing stroke length	21	a	a	a
Van Dantzig et al., 2013	1.188	sedetary behavior	86	a		a
Lee, 2014	1.05	medicine adherence	12	b	a	a
Brumby et al., 2011	1.0 (task priority x performance, no choice), and 2.77 (task priority x performance, choice)	Information processing while using car simulation	24	a	a,c	a
Oshima, 2013	0.953	weight loss	56	b	a	b
Wang et al., 2013	0.928	body massages, stretching in computer use	39	b	a	a
Bentley et al., 2013	0.887	self-understanding in health behavior, wellbeing	60	f, b	b	b
Tulusan et al., 2013	0.835	driving eco-friendly	50	b	a	a
Qian et al., 2011	0.603	walking pace	20	a	a	a
Maltz et al., 2004	0.556 (distance), 0.317 (modality)	keeping distance to car in front	120 / 15 ***	a		a
Pellegrini et al., 2012	0.5198 for body weight	weight loss	51	a		b
Liu, 2014	0.471	time not working, stress	30	a	a	a
Spring et al., 2013	0.43	weight loss	70	h	d	b
Donmez et al., 2008	0.4268	braking, accelerating, glancing in driving in simulator	48	a	c	b
Bech-larsen et al., 2012	0.381	fruit and vegetable consumption	256	a		b
Willis et al., 2010	0.332 length, 0.451 volume	water usage, shower length	49 *	b	a	a
Ford et al., 2010	0.293	eating behavior in obese children	106	a		b
Schulz et al., 2014	0.28 (t1, sequential) and 0.18 (t2, simultaneous)	health behavior	5055	a, d	d	b
Ahlamad et al., 2012	0.143	Home energy use	151	b	d	a

Kim et al., 2008	0.107	knowledge of peers' sleeping behavior	6	b	a, b	a
Quintal et al., 2012	0.052	electricity consumption	13 *	e	a, b	a

1 – Analysis method: a = Analysis of Variance, b = T-test, c = Nonparametric tests (e.g. Wilcoxon Signed Ranks), d = (Pearson's) Chi squared test, e = Correlations and regression, f = Descriptives only, h = Other

2 – Problems: a = underpowered, b = no control condition, c = lacking conditions, d = other (such as self-report measures, self-selection, sample distribution issues)

3 – Field: a = design research, hci, engineering; b = health and psychology

: number of households included in study; ** number of classes included in study; * experimental condition / control condition*

3.3 The effect of feedback through digital technology on *durable* habit change

The durability of the hypothesized effect was tested in only four of the 72 studies, three of which found at least partial evidence of lasting effects. A combination of a standard behavioral weight loss protocol and feedback from digital technology led to lasting weight loss after half a year of use (Pellegrini et al., 2012); a range of lifestyle-oriented interventions based on feedback had effects that were discernable even after two years after the single point intervention (Schultz et al., 2011); and delivering feedback to reduce eating rate led to a lasting decrease in weight after a year of use, which was still discernable six months after intervention completion (Ford et al., 2010).

Contrarily, in a study of thirteen households that involved an in-home display of energy use, Quintal, Pereira and Nunes (2012) found no significant effects of display use on energy consumption even after a full year. However, this lack of findings may be due to a lack of control conditions and/or low statistical power, since descriptive data do point in the direction of a positive effect.

For behavior change to take effect, however, sustained use of the intervention is needed: intervention adherence is known to be significantly correlated with intervention success (Burke et al, 2008). Only three papers looked into sustained use of the feedback technology. First, in a qualitative study on the use of health mash-ups translating information from different feedback sources into natural language, almost all participants used the intervention for the full 90 days of the project (Bentley et al., 2013). Contrarily, in a weight loss intervention (Pellegrini et al., 2012), 20% of participants stopped within 6 months; and Pereira, Quintal, Nunes, and Bergés (2012) found that even though they could report initial success, after four weeks interest in their feedback intervention on energy use was waning, with detrimental results on feedback effect. These latter two findings are in line with literature on sustained use of behavior change interventions, which show a sharp decline in self-monitoring willingness after 10-14 days (e.g. Burke et al., 2008; Patrick et al., 2009) and a linear decline of the use of wearable technology which results in about 40% dropout within 12 months (Ledger & McCaffrey, 2014).

3.4 The effect of interpersonal and intrapersonal differences

Previous research has shown that not everybody benefits equally from feedback interventions. Both stable (traits) and dynamic (states) moderators are seen to influence feedback efficacy. Surprisingly, only one study in this review looked directly at the effect of demographic variables on feedback effectiveness. In an analysis of feedback on energy use in 2000 households, Vassileva et al. (2012) found that socio-economic factors such as income, age and type of housing interacted with the preferred medium of feedback delivery. Unfortunately, their work did not include the effect of socio-economic status on feedback effect.

In a similar vein, only a few papers took individual differences of any kind into account, be it motivation, self-regulatory capabilities, or personality traits. Bech-Larsen & Grønhøj (2013) found that people who consumed hardly any fruit benefited more from feedback than people who already consumed close to the desired target, suggesting a ceiling effect to feedback effectiveness that would cause underperformers to benefit more from feedback interventions than high performers. Similarly, Tasic et al. (2012) found that people who used a lot of water for showering decreased their water use a lot more than people who used less. Wallenborn et al. (2011) found that men were more interested in the use of smart meters than women and indeed used them more.

Finally, the null result in research reported by Alahmad et al. (2012) might be seen as a further indication of ceiling effects in feedback interventions. If self-selection has a detrimental effect on the effectiveness of a feedback intervention, it might be that this is because participants are already performing the desired behavior to the maximum possible extent.

3.5 The effect of feedback technology and properties

Feedback content factors (such as feedback sign, level of tailoring, and comparison level), the technology through which the feedback is delivered, feedback characteristics (such as timing, modality, frequency and duration), and the presentation of the feedback, all may influence the efficacy of feedback interventions. In this section, we first present results regarding feedback content, followed by results regarding feedback technology, characteristics and design.

For each study, we analyzed the sign of the feedback, i.e. whether the digital technology delivered positive feedback ("You have exceeded your goal by 1,000 steps"), negative feedback ("you are still 1,000 steps short of your goal") or neutral feedback ("you have managed 9,000 steps today"). Furthermore, we analyzed the comparisons the digital technology made in delivering the data, i.e. comparing to past performance, peer behavior, or abstract norms. Level of tailoring was not taken into account, because every study in the review included some form of tailoring.

Feedback sign. The vast majority of studies (55 out of 72) delivered feedback in such a way that both positive and negative feedback were possible, 4 studies only utilized feedback with a negative sign, and two studies only provided positive feedback. A further 12 studies provided neutral feedback, i.e. without any form of reference to performance goals or norms and therefore without sign. Two of these twelve studies combined neutral feedback for one dependent variable with signed feedback for another dependent variable. In one study, the feedback was described without detail, so no feedback sign could be established.

Only two studies directly compared positive and negative feedback. Both studies, which compared the effect of rewards and penalties on engagement (Jain, Taylor, & Pescheira, 2012), and the effect of positive with negative feedback on work pace interruptions (Liu & Pfaff, 2014), found a greater effect for positive feedback than negative. Moreover, the latter study found that negative feedback does indeed increase performance, but at the cost of a greater stress level.

Feedback Comparison. Different forms of comparisons can be made with feedback data. Current performance can be compared to past performance (historic comparison), a social comparison with peers or unknown counterparts can be delivered, or performance can be compared to a norm or a goal (normative comparison). In this review, 52 studies made a normative comparison in their feedback. 18 studies gave historic comparisons (8 of which combining this with normative feedback, 1 with social feedback, and 2 with normative and social feedback), 7 studies used social comparison (3 of which in combination with other forms of comparisons). 7 studies delivered the data 'as is', without comparison. One study described the feedback without detail, so no information about comparison could be extracted.

Two studies contrasted different kinds of comparisons directly. Jain, Taylor, and Pescheira (2012) looked at the effect of normative and historic feedback comparisons in smart energy meters, finding that historic comparisons resulted in greater effect, whereas normative comparisons did not change energy use. In contrast, Sundramoorthy et al. (2011) found that normative, social and historic comparisons resulted in greater energy saving.

All in all, on the basis of the data extracted in this review, it is not possible to ascribe a more positive effect on feedback efficacy to a single strategy of comparison. This reflects findings in literature on feedback in general.

Feedback technology. To deliver the feedback, 16 studies utilized a mobile phone app, 11 studies used an in-home display – mostly for energy use monitoring –, in 9 studies feedback was delivered using a website, and 7 studies used a computer or tablet application. Four studies provided participants with a wearable device capable of delivering vibrotactile feedback and three studies used a driving simulator. SMS text messaging, Facebook apps, and interactive public displays were used once. One study provided feedback both through a mobile phone app and a website. The largest category is that of the 'smart' devices, used in 18 of the studies. These devices often resemble generic household instruments, such as cutlery or scales, augmented with sensors and actuators. All but three studies derived the data for the feedback directly from the target behavior; three studies relied on self-report for the generation of feedback.

Each feedback technology has particular characteristics that impact the overall experience of the user. The wearable vibrotactile devices could only deliver feedback in their own modality, concurrent with behavior, and without possibilities for comparison to earlier results or performance of others. SMS text messages could only be delivered retrospectively, as they rely on technology with a time lag. However, technology choice was not associated with differences in effects on habit disruption or change; positive results as well as null findings were spread evenly across technologies. Unfortunately, none of the studies in the analysis directly compared different technological channels.

Feedback timing. Of the reviewed studies, 20 delivered retrospective feedback, i.e. feedback after the behavior had been performed. 52 studies delivered concurrent feedback, i.e. during behavior performance. Two studies offered both forms for different behaviors, without a direct comparison. One study (Donmez, Boyle, & Lee, 2007) directly compared the effectiveness of feedback timing on behavior. In this study, a combination of retrospective and concurrent feedback yields greater effect than separate timing strategies, because of the additional informational benefit offered by recurrent feedback on top of the direct intervention in behavior offered by concurrent feedback. Furthermore, Tulusan, Staake, and Fleisch (2012) find that users of their eco-driving support application prefer direct, concurrent feedback over retrospective feedback: the efficacy of the application is significantly predicted by the usage of the direct feedback delivered by the app, but not by retrospective, indirect feedback.

Feedback modality. Of the papers included in the review, 58 studies offered visual feedback only, one offered auditory feedback only, and 8 studies used tactile feedback only. Five studies directly compared the effectiveness of different feedback modalities, two of which contrasted visual with auditory feedback, one study contrasted auditory with tactile feedback; one study contrasted visual with tactile feedback, and one study compared three feedback modes: visual, auditory and tactile. Studies comparing tactile feedback with other modalities found this modality more effective when aimed at changing motor skills (Maltz & Shinar, 2004; Epstein et al., 2012) and when disruptiveness mattered. Generally, tactile feedback was found to be less disruptive in other tasks compared to auditory feedback, which in turn is less disruptive than visual feedback. A reverse pattern can be observed in the amount of detail that can be communicated through different feedback modalities: visual feedback can be more detailed than auditory, which can offer more detail than tactile feedback (Hoggan & Brewster, 2010). One study (Epstein et al., 2012) reported an effect of feedback modality on the durability of the achieved behavior change: sitting posture was changed beneficially through visual feedback, but only the addition of tactile feedback on optimal posture led to lasting effects.

These studies serve as an indication that the optimal selection of feedback modality not only depends on the targeted behavior, but also on the amount of disruption that a given task allows and the necessary detail of the feedback. More evidence to support this assumption is needed.

Three papers support the assumption that multimodal feedback is more effective than single-mode feedback (Hoggan & Brewster, 2010; Lieberman & Breazeal, 2007; Quian et al., 2011). In these cases, the increased effect mostly lies in additional strengths of different feedback mode, for example tactile feedback in smartphones being more effective in noisy areas and auditory feedback more effective in silent areas. Maltz and Shinar (2004) tested the concurrent application of visual and auditory feedback in driving behavior and found no beneficial effect of multimodal feedback, leading to the conclusion that auditory feedback is most effective for driving behaviors and other modalities do not add further improvement.

Feedback frequency and duration. The greater part of included studies (67 out of 72) used either continuous or on-demand delivery of feedback, which means almost all studies made use of the possibilities digital technologies offer in delivering the feedback as soon as possible. No studies compared the effect of different delivery frequencies directly. From the current literature, no conclusions can be drawn on the effectiveness of feedback frequency on feedback impact.

The duration of the feedback intervention differed from a single trial to one year. Those papers reporting lasting intervention effects had durations of six months (Pellegrini et al., 2012; Schultz et al., 2014), and one year (Ford et al., 2010). However, there is an obvious confound of intervention length with the type of behavior targeted, because not every habitual behavior is equally difficult to change, with periods needed for change ranging from a few weeks to behavioral vigilance without time limit (Lally & Gardner, 2013). Therefore, a single standard of ideal feedback intervention duration and frequency seems conceptually impossible.

Feedback presentation: Usability and aesthetics. Three papers considered the effect of visual design on feedback effectiveness directly. All three found some explorative indication that design and aesthetics matter for feedback acceptance, use of the feedback device and feedback impact. One paper (Consolvo, MacDonald & Landay, 2009) provides a very useful list of directives for the design of feedback presentation. The authors state that feedback should be abstract and reflective, unobtrusive and public, aesthetically pleasing, positive, controllable, trending/historical in comparison, and comprehensive. Two papers (Nakajima & Lehdonvirta, 2011); Rodgers & Bartram, 2011) described how heightened abstraction and aesthetic pleasingness seem to come at a cost in terms of usability and comprehension.

3.6 Other insights

Close scrutiny of all reviewed studies revealed a couple of noteworthy additional themes that were not detected in the analysis of existing literature that led to the hypotheses posed in this review.

One additional theme that emerged is the role of disruption in feedback efficacy. Feedback can play a role in habit change by disrupting the automatic response to a cue.

However, this disruption may also cause a task to be abandoned or otherwise disturb task resumption (Bittner & Zondervan, 2015). The amount of disruption therefore needs to be carefully tailored to break the automatic cue-response-chain without abandoning the task altogether. In this analysis, two papers mentioned the role of disruptiveness on feedback effect. As mentioned above in the section on feedback modality, a study of feedback delivered by a mobile game with different feedback modalities (Hoggan, Crossan, Brewster, & Kaaresoja, 2009) exhibited an interaction between feedback modality, disruption, and richness of the feedback. Interestingly, one study (Liu & Pfaff, 2014) showed how feedback can also be used to facilitate the resumption of tasks after disruptions.

Another important insight is that the amount of integration of feedback in other areas of behavior, such as usage of similar interventions or sharing behavior on online social networks, might be a strong predictor of feedback effect. Wallenborn, Orsini and Vanhaverbeeke (2011) found that when energy monitors are not integrated in pre-existing practices, the information quickly disappears into background noise like with any other new appliance. A study by Jain, Taylor & Pescheira (2012) had a similar finding in a study of the usage of an interface providing feedback on energy consumption. Bentley et al. (2013) found similar patterns in the effect of health mashups. When participants used an app that integrated fitbit activity tracking data with weight, food intake, sleep etcetera, sustained use of the feedback technology increased.

This notion of integration is an interesting concept that needs further exploration. Indeed, relevant theories that explain the effectiveness of feedback on behavior change, such as Social Cognitive Theory (e.g. Bandura, 1997) or Control Theory (Kuhl, 1985; Carver & Scheier, 1985), suggest that behavior change is most likely if feedback is not delivered on its own, but embedded in larger interventions with clear target behaviors and action plans. This notion is also backed up by considerable evidence from original research (e.g. Avery et al., 2012; Sniehotta et al, 2006; Godino et al., 2013) and reviews (e.g. Dombrowski et al., 2012; Gardner et al., 2010).

Wallenborn, Orsini and Vanhaverbeeke (2011) noted that wasteful behavior in energy use can arise from role perception ("a good parent always gets the laundry clean and therefore washes at 90˚ C") and different levels of technical insight in families might lead to conflicts about the performance on feedback. This gives insight in how social interactions influence feedback effect. Feedback on performance spurs discussion with family members and others, which may in itself lead to behavior change or even conflicts and role clashes. Similar effects are reported by Kappel and Grechenig (2009) when they mention positive social effects of their device that reports water usage in the shower: "A couple used to argue that one of them always took longer in the shower and (...) used more water. (...) (T)hey learned that the woman used only half as much water, even though she spent more time in the shower. This discovery stimulated the man to further reduce his own water consumption. In another household the child (11 yrs.) triggered discussions about the water consumption, because he used much less water than his parents. This stimulated his mother to begin reducing her own consumption (...)." Nakajima & Lehdonvirta (2013) and Katzeff et al. (2013) found similar results in an intervention aimed at (respectively) children's tooth brushing and energy use in the office.

4. Conclusion

This review shows that in the 72 studies we analyzed, feedback delivered through digital technology is generally effective in disrupting habitual behavior. However, the current literature does not provide enough evidence to support the hypothesis that feedback through digital technology leads to lasting behavior change. Furthermore, little is known about factors that facilitate sustained use of digital technology, intra-personal and inter-personal moderators of feedback efficacy, and the effect of feedback characteristics.

This review makes clear that feedback through digital technology has the potential to disrupt undesired habits. Therefore, such feedback can be seen as a potentially reinforcing ingredient for any intervention aimed at habit change. This work offers support for Quantified Self-solutions, which may indeed lead to healthier, more eco-friendly behaviors; it also supports the notion that delivering feedback through digital technology may heighten the chances of conscious scrutiny for a broad range of deeply engrained, undesirable habits. Our analysis shows this finding is consistent across feedback technologies: feedback delivered through a broad range of technological channels appears to succeed in disrupting undesired habits.

However, the possibilities of using feedback through digital technology for sustainable habit change have yet to be proven. Particularly, the durability of the feedback effect on habitual behavior is as yet unclear. Those few studies that included longitudinal measurements generally found sustainable effects of feedback on behavior, but the greater part of the studies only measured effects right after the intervention. To prove the hypothesis that feedback through digital technology actually enables users to change their behavior, more evidence on whether the use of the digital technology leads to lasting effects is necessary.

To ensure the occurrence of behavior change, intervention designers must make sure their technology is accepted by its users, and used long enough to warrant habit change. Existing literature (e.g. Ledger & McCaffrey, 2014) suggests that technological feedback solutions are often to be discarded after initial use. Unfortunately, methods to maintain engagement with a technology over time remain unclear.

The role of moderating traits and demographic factors also remains understudied. Very little is known of the interplay of traits and states on the one hand, and feedback properties such as feedback sign, comparison, and delivery mode on the other. Similarly, the effect of different feedback properties such as timing, modality, frequency and duration, have not yet received the attention needed to draw any conclusions on their impact on feedback effect. This suggests that we cannot yet tell whether changes in behavior can really be attributed to the digital technology and its feedback, or that these are merely functioning as some sort of lens through which only well-motivated and capable individuals manage to focus their behavior-changing endeavors.

Although this review provides evidence for the effect of feedback through digital technology on disrupting habitual behavior, this review also demonstrates that research into such effects has only just started. Because of the explorative, descriptive nature of a great part of the included papers, there are limits to the conclusions that can be drawn from this review. The majority of the included quantitative studies, 33 out of 50, did not report any control measures for demand characteristics or other experimental artifacts, e.g. through well-balanced experimental designs. Furthermore, 22 out of 30 quantitative studies with statistical analysis were statistically underpowered, which seriously reduces the validity of any conclusions drawn from those papers. As a consequence, only a part of the 72 original studies in this review (14 quantitative studies and 21 qualitative studies) were described in a way that proves enough methodological rigor to act as a source for direct evidence. The literature would benefit greatly from well-performed additional research on the effect of feedback through digital technology on habitual behavior, be it field studies or lab work, with good active controls for experimental artifacts and ample statistical power.

Moreover, it remains unknown how many studies did not make the literature because the desired effect could not be shown or no support was found for the original hypothesis. The great majority of studies in this review found a positive effect of feedback on habit disruption, much more so than in similar analyses (e.g. Kluger and Denisi, 1996, who find a 66% success rate). The field (and science in general) would greatly benefit from measures aimed at reducing publication bias, such as pre-registering studies, to provide insight into how many 'failed' studies end up in the proverbial file drawer (Franco, Malhotra, & Simonovits, 2014).

The review also shows the merit of combining quantitative research with good qualitative and explorative research. It is paramount that theories of behavior change are supported by well-designed trials, but important insights such as the influence of social interaction on the effects of feedback delivered by digital technology would not easily show up in even the most well-set up quantitative research.

4.1 Further research

All of these areas provide ample possibilities for further research. The broad range of dependent variables and feedback technologies limit the validity and generalizability of the findings in this review. However, the results presented here may serve as a basis for further studies and analyses.

One such analysis could examine which behaviors are most likely to benefit from feedback delivered through digital technology. Intuitively, the hypothesis that feedback does not affect every habitual behavior equally seems plausible, but evidence is lacking. Similar questions arise when the different technologies are taken into view. Different technologies offer different possibilities for feedback modality and other properties. It seems plausible to assume that these differences influence efficacy, but this does not follow from the results of this review. Particular attention should be paid to the level of disruption of the feedback. Evidence (Bittner & Zondervan, 2015) suggests that feedback may disrupt tasks in such a way that this leads to task abandonment. Some feedback modalities (visual) are clearly more disruptive than others (vibrotactile, auditive). The effects of feedback disruptiveness on sustained performance warrant further scrutiny.

Factors moderating the sustained use of technological solutions are another area that deserves our attention. Without use, we cannot expect technology to have any effect on behavior. User experience, usability, and design can be thought of as moderating factors on the effect of feedback, but as yet this hypothesis lacks support. Intuitively, and from what little evidence that exists (e.g. Ludden et al, 2015), one would reason that clunky designs are unlikely to get used, with detrimental consequences. Therefore, we see the lack of focus on usability in this research field as a serious problem. Similar focus is needed on other factors influencing the lasting use of technological feedback solutions. Is a high motivation essential? Do certain personality characteristics facilitate sustained use, and what is the effect of feedback characteristics? All these questions need an answer.

Another example of an area of interest that deserves further scrutiny is the effect of personality traits and states such as initial motivation and self-efficacy on feedback impact. Literature suggests that high initial motivation and self-efficacy increase the impact of feedback on habitual behavior. However, results from studies in this paper suggest a ceiling effect. A well-set up experimental design could shed light on the effect of initial motivation and perceived self-efficacy on the effect of feedback on habits.

A similar question remains about the effect of feedback sign. In this review, the greater part of the studies provided feedback in such a way that both positive and negative feedback was possible. Unfortunately, this makes it impossible to test an interesting hypothesis, i.e. concerning the interaction between feedback sign and regulatory focus – the tendency to approach positive impulses and avoid negative ones. Van Dijk and Kluger (1994, 2011) suggest that in a prevention focus (avoiding negative consequences), negative feedback should have more effect, whilst in a promotion focus (approaching positive consequences), positive feedback should have more effect. Hattie and Timperley (2007) however, find in a meta-analysis that positive feedback should always lead to more effect than negative feedback. This issue is particularly relevant to feedback delivered through digital technology, which by nature is capable of delivering both signs, depending on individual performance. Is feedback more effective in a prevention focus as long as goals are being reached, and does it lose its effect when goals are too hard - and similarly, is feedback more effective in a prevention focus as long as goals are not reached yet? Further research could give valuable insights in when feedback through digital technology has the most effect.

In a similar vein, the optimal choice of feedback properties in such a way that feedback is delivered concurrently with behavior in a continuous or on-demand manner, and data gathering for the feedback takes place automatically without the need for self report measures, should intuitively lead to an enhanced feedback efficacy. This hypothesis, however, remains unsubstantiated. Subjects of similar interest that have not been researched in a controlled manner at all are the active integration of feedback through digital technology within more complex interventions, and the social effects of digital technology. In real-life situations, feedback is not delivered in a vacuum, but plays a role in a social practice. Users will interact with friends, family and others about the received feedback, the attainability of goals, and the use of the artifact that delivers the feedback. The effects of feedback integration and social practices on feedback efficacy are in urgent need of research.

Further research into the effectiveness of feedback interventions to disrupt habits, personal differences in feedback efficacy, and the effect of applying different feedback characteristics, might not only enhance our knowledge on how habits might be changed. Such research would also serve as a basis for intervention developers and designers to inform the design of more effective behavior change products. The ubiquity of Quantified Self-solutions and health-related apps on smartphones show a great level of acceptance of this kind of intervention. The public is generally ready and willing to embrace such interventions. Badly set-up products without a base in scientific evidence might do lasting damage to the benevolent reception feedback interventions currently receive. But well-designed, evidence-based solutions can be expected to have a great impact on our well-being and on the proliferation of sustainable behavior. Feedback through digital technology as an intervention strategy to change undesirable habitual behavior offers great chances for healthier and more sustainable living that should not be wasted.

References

Aarts, H., & Dijksterhuis, A. (2003). The silence of the library: environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, *84*(1), 18.

* Alahmad, M. A., Wheeler, P. G., Schwer, A., Eiden, J., & Brumbaugh, A. (2012). A comparative study of three feedback devices for residential real-time energy monitoring. *Industrial Electronics, IEEE Transactions on*, *59*(4), 2002-2013.

* Alahakone, A. U., & Senanayake, S. A. (2009). A Real Time Vibrotactile Biofeedback System for Improving Lower Extremity Kinematic Motion during Sports Training. In *Soft Computing and Pattern Recognition, 2009. SOCPAR'09. International Conference of* (pp. 610-615). IEEE.

* Alahakone, A. U., & Senanayake, S. A. (2010). A real-time system with assistive feedback for postural control in rehabilitation. *Mechatronics, IEEE/ASME Transactions on*, *15*(2), 226-233.

Alberts, H. J., Martijn, C., & de Vries, N. K. (2011). Fighting self-control failure: Overcoming ego depletion by increasing self-awareness. *Journal of Experimental Social Psychology*, *47*(1), 58-62.

Andler, B., Woolf, S. & Wilson, T. (2013). Uji shower head. Product under development. Retrieved 09/25, 2013, from http://ujishower.com/.

Avery, A. J., Rodgers, S., Cantrill, J. A., Armstrong, S., Cresswell, K., Eden, M., ... & Sheikh, A. (2012). A pharmacist-led information technology intervention for medication errors (PINCER): a multicentre, cluster randomised, controlled trial and cost-effectiveness analysis. *The Lancet*, *379*(9823), 1310-1319.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control.* New York, NY: Freeman.

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues. *Handbook of Social Cognition*, *1*, 1-40. Oxford: Psychology Press.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource?. *Journal of Personality and Social Psychology*, *74*(5), 1252.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*(1), 1-15.

* Bech-Larsen, T., & Grønhøj, A. (2013). Promoting healthy eating to children: a text message (SMS) feedback approach. *International Journal of Consumer Studies*, *37*(3), 250-256.

* Bentley, F., Tollmar, K., Stephenson, P., Levy, L., Jones, B., Robertson, S., ... & Wilson, J. (2013). Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *20*(5), 30.

Bittner, J. V., & Zondervan, R. (2015). Motivating and achievement-eliciting pop-ups in online environments: A user experience perspective. *Computers in Human Behavior*, *50*, 449-455.

Borenstein, M., Hedges, L.V., Higgins, J.P., Rothstein, H.R. (2009). *Introduction to Meta-analysis.* New York, NY: Wiley & Sons.

Burke, L. E., Sereika, S. M., Music, E., Warziski, M., Styn, M. A., & Stone, A. (2008). Using instrumented paper diaries to document self-monitoring patterns in weight loss. *Contemporary Clinical Trials*, *29*(2), 182-193.

Braverman, J. (2008). Testimonials Versus Informational Persuasive Messages: The Moderating Effect of Delivery Mode and Personal Involvement. *Communication Research*, *35*(5), 666-694.

* Brumby, D. P., Davies, S. C., Janssen, C. P., & Grace, J. J. (2011). Fast or safe?: how performance objectives determine modality output choices while interacting on the move. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 473-482). ACM.

* Bruns Alonso, M., Stienstra, J., & Dijkstra, R. (2014). Brush and learn: transforming tooth brushing behavior through interactive materiality, a design exploration. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction* (pp. 113-120). ACM.

Brug, J., Glanz, K., Van Assema, P., Kok, G., & Van Breukelen, G. J. (1998). The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. *Health Education & Behavior*, *25*(4), 517-531.

Carver, C. S., & Scheier, M. F. (1985). A control-systems approach to the self-regulation of action. In: Kuhl, J. *Action Control* (pp. 237-265). Berlin: Springer Verlag.

Carver, C.S., & Scheier, M.F. (2011). Self-Regulation of Action and Affect. in: Vohs, K. D., & Baumeister, R. F. (Eds.). *Handbook of Self-Regulation: Research, Theory, and Applications*. New York, NY: Guilford Press.

* Chang, Y. C., Lo, J. L., Huang, C. J., Hsu, N. Y., Chu, H. H., Wang, H. Y., ... & Hsieh, Y. L. (2008). Playful toothbrush: ubicomp technology for teaching tooth brushing to kindergarten children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 363-372). ACM.

* Comber, R., & Thieme, A. (2013). Designing beyond habit: opening space for improved recycling and food waste behaviors through processes of persuasion, social influence and aversive affect. *Personal and Ubiquitous Computing*, *17*(6), 1197-1210.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

* Connelly, K. H., Faber, A. M., Rogers, Y., Siek, K. A., & Toscos, T. (2006). Mobile applications that empower people to monitor their personal health. *e & i Elektrotechnik und Informationstechnik*, *123*(4), 124-128.

* Consolvo, S., McDonald, D. W., & Landay, J. A. (2009). Theory-driven design strategies for technologies that support behavior change in everyday life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 405-414). ACM.

* Cowan, B. R., Bowers, C. P., Beale, R., & Pinder, C. (2013). The stroppy kettle: an intervention to break energy consumption habits. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1485-1490). ACM.

Darby, S. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, *486*, 2006.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, *125*(6), 627.

Donovan, J. J., & Hafsteinsson, L. G. (2006). The Impact of Goal-Performance Discrepancies, Self-Efficacy, and Goal Orientation on Upward Goal Revision1. *Journal of Applied Social Psychology*, *36*(4), 1046-1069.

Dombrowski, S.U., Sniehotta, F.F., Avenell, A., Johnston, M., MacLennan, G., & Araújo-Soares, V. (2012). Identifying active ingredients in complex behavioral interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review, *Health Psychology Review, 6,* 7-32,

* Donmez, B., Boyle, L. N., & Lee, J. D. (2008). Mitigating driver distraction with retrospective and concurrent feedback. *Accident Analysis & Prevention*, *40*(2), 776-786.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*(2), 73-92.

* Epstein, R., Colford, S., Epstein, E., Loye, B., & Walsh, M. (2012). The effects of feedback on computer workstation posture habits. *Work: A Journal of Prevention, Assessment and Rehabilitation*, *41*(1), 73-79.

* Erickson, T., Li, M., Kim, Y., Deshpande, A., Sahu, S., Chao, T., ... & Naphade, M. (2013). The dubuque electricity portal: evaluation of a city-scale residential electricity consumption feedback system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1203-1212). ACM.

Faruqui, A., Harris, D., & Hledik, R. (2010). Unlocking the € 53 billion savings from smart meters in the EU: How increasing the adoption of dynamic tariffs could make or break the EU's smart grid investment. *Energy Policy*, *38*(10), 6222-6231.

Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy?. *Energy Efficiency,* 1(1), 79-104.

Fogg, B.J. (2003). *Persuasive Technology; Using Computers to Change what we think and do.* San Francisco, CA: Morgan Kaufmann.

* Ford, A. L., Bergh, C., Södersten, P., Sabin, M. A., Hollinghurst, S., Hunt, L. P., & Shield, J. P. (2010). Treatment of childhood obesity by retraining eating behavior: randomised controlled trial. *BMJ*, *340*.

Fox, S., & Duggan, M. (2012). *Mobile Health 2012.* Washington DC: Pew Research Center's Internet & American Life Project, 2012. Retrieved 7/15/2014 from www.pewinternet.org/Reports/2012/Mobile-Health.aspx.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502-1505.

* Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., & Landay, J. A. (2009). UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1043-1052). ACM.

Froehlich, J., Findlater, L., & Landay, J. (2010). The design of eco-feedback technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1999-2008). ACM.

* Ganglbauer, E., Fitzpatrick, G., & Comber, R. (2013). Negotiating food waste: Using a practice lens to inform design. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *20*(2), 11.

Gardner, B. (2014). A Review and Analysis of the Use of 'Habit' in Understanding, Predicting and Influencing Health-related Behavior. *Health Psychology Review, 8*(1), 1-19.

Gardner, B., Whittington, C., McAteer, J., Eccles, M. P., & Michie, S. (2010). Using theory to synthesise evidence from behavior change interventions: the example of audit and feedback. *Social Science & Medicine*, *70*(10), 1618-1625.

Garg, A. X., Hackam, D., & Tonelli, M. (2008). Systematic review and meta-analysis: when one study is just not enough. *Clinical Journal of the American Society of Nephrology*, *3*(1), 253-260.

Geister, S., Konradt, U., & Hertel, G. (2006). Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small Group Research*, *37*(5), 459-489.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing Type S (sign) and Type M (magnitude) errors1. *Perspectives on Psychological Science*, *1*, 11.

Godino, J. G., Watkinson, C., Corder, K., Marteau, T. M., Sutton, S., Sharp, S. J., ... & van Sluijs, E. M. (2013). Impact of Personalised Feedback about Physical Activity on Change in Objectively Measured Physical Activity (the FAB Study): A Randomised Controlled Trial. *PloS one*, *8*(9), e75398.

Guadagno, R. E., & Cialdini, R. B. (2007). Persuade him by email, but see her in person: Online persuasion revisited. *Computers in Human Behavior*, *23*(2), 999-1015.

* Gustafsson, A., Bång, M., & Svahn, M. (2009). Power explorer: a casual game style for encouraging long term behavior change among teenagers. In *Proceedings of the International Conference on Advances in Computer Enterntainment Technology* (pp. 182-189). ACM.

* Gyllensward, M., Gustafsson, A., & Bång, M. (2006). Visualizing energy consumption of radiators. In *Persuasive Technology* (pp. 167-170). Springer Berlin Heidelberg.

Hall, A. K., Cole-Lewis, H., & Bernhardt, J. M. (2015). Mobile text messaging for health: a systematic review of reviews. *Annual review of public health*, *36*, 393.

Hamilton, L. C. (1985). Self-reported and actual savings in a water conservation campaign. *Environment and Behavior*, *17*(3), 315-326.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

Heller, F., & Borchers, J. (2011). PowerSocket: towards on-outlet power consumption visualization. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1981-1986). ACM.

* Hargreaves, T., Nye, M., & Burgess, J. (2013). Keeping energy visible? Exploring how householders interact with feedback from smart energy monitors in the longer term. *Energy Policy*, *52*, 126-134.

Higgins, E. T. (1997). Beyond pleasure and pain. *American psychologist*, *52*(12), 1280.

Ho, V., Simmons, R. K., Ridgway, C. L., van Sluijs, E. M., Bamber, D. J., Goodyer, I. M., ... & Corder, K. (2013). Is wearing a pedometer associated with higher physical activity among adolescents?. *Preventive Medicine*, *56*(5), 273-277.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in cognitive sciences*, *16*(3), 174-180.

* Hoggan, E., & Brewster, S. A. (2010). Crosstrainer: testing the use of multimodal interfaces in situ. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 333-342). ACM.

Hoggan, E., Crossan, A., Brewster, S. A., & Kaaresoja, T. (2009). Audio or tactile feedback: which modality when?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2253-2256). ACM.

* Hurling, R., Catt, M., Boni, M. D., Fairley, B. W., Hurst, T., Murray, P., ... & Sodhi, J. S. (2007). Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. *Journal of Medical Internet Research*, *9*(2), e7.

Ivers, N., Jamtvedt, G., Flottorp, S., Young, J. M., Odgaard-Jensen, J., French, S. D., ... & Oxman, A. D. (2012). Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database System Review*, *6*.

Jager, W. (2003). Breaking 'bad habits': a dynamical perspective on habit formation and change. *Human Decision-Making and Environmental Perception–Understanding and Assisting Human Decision-Making in Real Life Settings. Libor Amicorum for Charles Vlek,* Groningen: University of Groningen.

* Jain, R. K., Taylor, J. E., & Peschiera, G. (2012). Assessing eco-feedback interface usage and design to drive energy efficiency in buildings. *Energy and Buildings*, *48*, 8-17.

* Kalnikaite, V., Rogers, Y., Bird, J., Villar, N., Bachour, K., Payne, S., ... & Kreitmayer, S. (2011). How to nudge in Situ: designing lambent devices to deliver salient information in supermarkets. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (pp. 11-20). ACM.

* Kappel, K., & Grechenig, T. (2009). Show-me: water consumption at a glance to promote water conservation in the shower. In *Proceedings of the 4th International Conference on Persuasive Technology* (p. 26). ACM.

Kaptein, M., De Ruyter, B., Markopoulos, P., & Aarts, E. (2012). Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *2*(2), 10.

* Katzeff, C., Broms, L., Jönsson, L., Westholm, U., & Räsänen, M. (2013). Exploring sustainable practices in workplace settings through visualizing electricity consumption. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *20*(5), 31.

* Kim, S., Kientz, J. A., Patel, S. N., & Abowd, G. D. (2008). Are you sleeping?: sharing portrayed sleeping status within a social network. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 619-628). ACM.

Kim, J. Y., Lee, K. H., Kim, S. H., Kim, K. H., Kim, J. H., Han, J. S., ... & Bae, W. K. (2013). Needs analysis and development of a tailored mobile message program linked with electronic health records for weight reduction. *International Journal of Medical Informatics*, *82*(11), 1123-1132.

* Kirkham, R., Mellor, S., Green, D., Lin, J. S., Ladha, K., Ladha, C., ... & Plötz, T. (2013). The break-time barometer: an exploratory system forworkplace break-time social awareness. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 73-82). ACM.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254.

Kuhl, J. (1985). *Action Control: From Cognition to Behavior*. Berlin: Springer Verlag.

* Kuznetsov, S., & Paulos, E. (2010). UpStream: motivating water conservation with low-cost water flow sensing and persuasive displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1851-1860). ACM.

* Kwon, D. Y., & Gross, M. (2005). Combining Body Sensors and Visual Sensors for Motion Tracking. *Proc. Advances in Computer Entertainment Technology,* (pp. 94–101). ACM.

Lally, P., & Gardner, B. (2013). Promoting habit formation. *Health Psychology Review*, *7*(sup1), 137-158.

Lam, C. F., DeRue, D. S., Karam, E. P., & Hollenbeck, J. R. (2011). The impact of feedback frequency on learning and task performance: Challenging the “more is better” assumption. *Organizational Behavior and Human Decision Processes*, *116*(2), 217-228.

* Lane, N. D., Lin, M., Mohammod, M., Yang, X., Lu, H., Cardone, G., ... & Choudhury, T. (2014). BeWell: Sensing Sleep, Physical Activities and Social Interactions to Promote Wellbeing. *Mobile Networks and Applications*, 1-15.

Larson, K., & Picard, R. (2005,). The Aesthetics of Reading. In *Human-Computer Interaction Consortium Conference, Snow Mountain Ranch, Fraser, Colorado*.

Laschke, M., Hassenzahl, M., & Diefenbach, S. (2011). Things with attitude: transformational products. In *Create11 Conference* (pp. 1-2). ACM.

* Lee, M. L., & Dey, A. K. (2014). Real-time feedback for improving medication taking. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 2259-2268). ACM.

Ledger, D., & McCaffrey, D. (2014). Inside Wearables. How the Science of Human Behavior Change Offers the Secret to Long-Term Engagement. Retrieved on 24 November 2014 from http://endeavourpartners.net/assets/Endeavour-Partners-Wearables-and-the-Science-of-Human-Behavior-Change-Part-1-January-20141.pdf

Li, I., Dey, A.K., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In: *Proceedings of CHI 2010,* 557–566.

Lichtman, S. W., Pisarska, K., Berman, E. R., Pestone, M., Dowling, H., Offenbacher, E., ... & Heymsfield, S. B. (1992). Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *New England Journal of Medicine*, *327*(27), 1893-1898.

* Lieberman, J., & Breazeal, C. (2007). TIKL: Development of a Wearable Vibrotactile Feedback Suit for Improved Human Motor Learning. *IEEE Transactions on Robotics*, 919–926.

* Liu, Y., & Pfaff, M. S. (2014). Wiredin: using visual feedback to support task resumption. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems* (pp. 2539-2544). ACM.

Ludden, G. D., van Rompay, T. J., Kelders, S. M., & van Gemert-Pijnen, J. E. (2015). How to Increase Reach and Adherence of Web-Based Interventions: A Design Research Viewpoint. *Journal of Medical Internet Research*, *17*(7).

Maitland, J., Chalmers, M., & Siek, K. A. (2009). Persuasion not required. Improving our understanding of the sociotechnical context of dietary behavioral change. In *Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009. 3rd International Conference on* (pp. 1-8). IEEE.

* Maltz, M., & Shinar, D. (2004). Imperfect in-vehicle collision avoidance warning systems can aid drivers. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(2), 357-366.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (Vol. 1). Psychology Press.

McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC Medical Research Methodology*, *7*(1), 30.

Millock, K., & Nauges, C. (2010). Household adoption of water-efficient equipment: the role of socio-economic factors, environmental attitudes and policy. *Environmental Resource Economics, 46*, 539–565.

Moynihan, T. (2015). Apple's ResearchKit Is a New Way to Do Medical Research | WIRED. Retrieved March 17, 2015, from http://www.wired.com/2015/03/apples-new-researchkit-app-framework-medical-research/

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counseling psychology*, *38*(1), 30.

* Nakajima, T., & Lehdonvirta, V. (2013). Designing motivation using persuasive ambient mirrors. *Personal and Ubiquitous Computing*, *17*(1), 107-126.

Nelson, R. O., & Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification*, *5*(1), 3-14.

* Nelson, T., Megens, C., & Peeters, M. (2012). Bouncers: a design exploration into sharing and triggering personal activity amongst friends. *Persuasive Technology*, 33.

Noar, S. M., Benac, C. N., & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673.

Noar, S. M., Harrington, N. G., Van Stee, S. K., & Aldrich, R. S. (2011). Tailored health communication to change lifestyle behaviors. *American Journal of Lifestyle Medicine*, *5*(2), 112-122.

Northcraft, G. B., Schmidt, A. M., & Ashford, S. J. (2011). Feedback and the rationing of time and effort among competing tasks. *Journal of Applied Psychology*, *96*(5), 1076.

O’Kane, A. A., Rogers, Y., & Blandford, A. (2015). Concealing or Revealing Mobile Medical Devices?: Designing for Onstage and Offstage Presentation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1689-1698). ACM.

Orne, M. T. (2009). Demand characteristics and the concept of quasi-controls. *Artifacts in Behavioral Research: Robert Rosenthal and Ralph L. Rosnow's Classic Books*, 110.

* Oshima, Y., Matsuoka, Y., & Sakane, N. (2013). Effect of weight-loss program using self-weighing twice a day and feedback in overweight and obese subject: A randomized controlled trial. *Obesity Research & Clinical Practice*, *7*(5), e361-e366.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. New York, NY: Sage Publications.

Patrick, K., Raab, F., Adams, M. A., Dillon, L., Zabinski, M., Rock, C. L., ... & Norman, G. J. (2009). A text message–based intervention for weight loss: randomized controlled trial. *Journal of medical Internet research*, *11*(1).

* Pellegrini, C. A., Verba, S. D., Otto, A. D., Helsel, D. L., Davis, K. K., & Jakicic, J. M. (2012). The Comparison of a Technology-Based System and an In-Person Behavioral Weight Loss Intervention. *Obesity*, *20*(2), 356-363.

* Pereira, L., Quintal, F., Nunes, N., & Bergés, M. (2012). The design of a hardware-software platform for long-term energy eco-feedback research. In *Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (pp. 221-230). ACM

* Pollak, J., Gay, G., Byrne, S., Wagner, E., Retelny, D., & Humphreys, L. (2010). It's time to eat! Using mobile games to promote healthy eating. *Pervasive Computing, IEEE*, *9*(3), 21-27.

Pursey, K., Burrows, T. L., Stanwell, P., & Collins, C. E. (2014). How Accurate is Web-Based Self-Reported Height, Weight, and Body Mass Index in Young Adults?. *Journal of Medical Internet Research*, *16*(1).

* Qian, H., Kuber, R., Sears, A., & Murphy, E. (2011). Maintaining and modifying pace through tactile and multimodal feedback. *Interacting with Computers*, *23*(3), 214-225.

Quinn, J. M., Pascoe, A., Wood, W., & Neal, D. T. (2010). Can't control yourself? Monitor those bad habits. *Personality and Social Psychology Bulletin*, *36*(4), 499-511.

* Quintal, F., Pereira, L., & Nunes, N. J. (2012). A long-term study of energy eco-feedback using non-intrusive load monitoring. *Persuasive Technology*, 49.

Quintana, D. S. (2015). From pre-registration to publication: a nontechnical primer for conducting a meta-analysis to synthesize correlational data. *Frontiers in Psychology*, *6*, 1549.

* Reitberger, W. H., Spreicer, W., & Fitzpatrick, G. (2014). Nutriflect: reflecting collective shopping behavior and nutrition. In *Proceedings of the 32nd annual ACM Conference on Human Factors in Computing Systems* (pp. 3309-3318). ACM.

* Rodgers, J., & Bartram, L. (2011). Exploring ambient and artistic visualization for residential energy use feedback. *Visualization and Computer Graphics, IEEE Transactions on*, *17*(12), 2489-2497.

* Rodríguez, M. D., Roa, J. R., Morán, A. L., & Nava-Muñoz, S. (2013). CAMMInA: a mobile ambient information system to motivate elders to exercise. *Personal and Ubiquitous Computing*, *17*(6), 1127-1134.

Schoffman, D. E., Turner-McGrievy, G., Jones, S. J., & Wilcox, S. (2013). Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? *Translational Behavioral Medicine*, *3*(3), 320-325.

Schön, D. A. (1988). From technical rationality to reflection-in-action. *Professional Judgment: A Reader in Clinical Decision Making*, 60-77.

* Schulz, D. N., Kremers, S. P., Vandelanotte, C., van Adrichem, M. J., Schneider, F., Candel, M. J., & de Vries, H. (2014). Effects of a Web-Based Tailored Multiple-Lifestyle Intervention for Adults: A Two-Year Randomized Controlled Trial Comparing Sequential and Simultaneous Delivery Modes. *Journal of Medical Internet Research*, *16*(1), e26.

Sniehotta, F. F., Nagy, G., Scholz, U., & Schwarzer, R. (2006). The role of action control in implementing intentions during the first weeks of behavior change. *British Journal of Social Psychology*, *45*(1), 87-106.

* Spagnolli, A., Corradi, N., Gamberini, L., Hoggan, E., Jacucci, G., Katzeff, C., ... & Jönsson, L. (2011). Eco-feedback on the go: Motivating energy awareness. *Computer*, *44*(5), 38-45.

* Spelmezan, D. (2012). An investigation into the use of tactile instructions in snowboarding. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 417-426). ACM.

* Spring, B., Duncan, J. M., Janke, E. A., Kozak, A. T., McFadden, H. G., DeMott, A., ... & Hedeker, D. (2013). Integrating technology into standard weight loss treatment: a randomized controlled trial. *JAMA Internal Medicine*, *173*(2), 105-111.

* Stamopoulos, S. F., Komninos, A., & Garofalakis, I. (2014). A Mobile Shopping Assistant to Support Product Domesticity in Consumer Decisions. In *Proceedings of the 18th Panhellenic Conference on Informatics* (pp. 1-6). ACM.

* Stienstra, J., Overbeeke, K., & Wensveen, S. (2011). Embodying complexity through movement sonification: case study on empowering the speed-skater. In *Proceedings of*

the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity (pp. 39-44). ACM.

* Strengers, Y. (2011). Negotiating everyday life: The role of energy and water consumption feedback. *Journal of Consumer Culture*, *11*(3), 319-338.

* Sundramoorthy, V., Cooper, G., Linge, N., & Liu, Q. (2011). Domesticating energy-monitoring systems: Challenges and design concerns. *IEEE Pervasive Computing*, *10*(1), 20-27.

* Tasic, V., Staake, T., Stiefmeier, T., Tiefenbeck, V., Fleisch, E., & Troster, G. (2012). Self-powered water meter for direct feedback. In *Internet of Things (IOT), 2012 3rd International Conference on the* (pp. 24-30). IEEE.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, *13*(2), 127-145.

Tuch, A. N., Presslaber, E. E., Stöcklin, M., Opwis, K., & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, *70*(11), 794-811.

* Tulusan, J., Staake, T., & Fleisch, E. (2012). Direct or indirect sensor enabled eco-driving feedback: Which preference do corporate car drivers have?. In *Internet of Things (IOT), 2012 3rd International Conference on the* (pp. 39-46). IEEE.

* Tulusan, J., Staake, T., & Fleisch, E. (2012). Providing eco-driving feedback to corporate car drivers: what impact does a smartphone application have on their fuel efficiency?. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp. 212-215). ACM.

* Ueno, T., Sano, F., Saeki, O., & Tsuji, K. (2006). Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. *Applied Energy, 83*(2), 166–183.

* Van Dantzig, S., Geleijnse, G., & Halteren, A. T. (2013). Toward a persuasive mobile application to reduce sedentary behavior. *Personal and Ubiquitous Computing*, *17*(6), 1237-1246.

* Van der Linden, J., Johnson, R., Bird, J., Rogers, Y., & Schoonderwaldt, E. (2011). Buzzing to play: lessons learned from an in the wild study of real-time vibrotactile feedback. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (pp. 533-542). ACM.

Van Dijk, D., & Kluger, A. N. (2004). Feedback sign effect on motivation: Is it moderated by regulatory focus? *Applied Psychology*, *53*(1), 113-135.

Van Dijk, D., & Kluger, A. N. (2011). Task type as a moderator of positive/negative feedback effects on motivation and performance: A regulatory focus perspective. *Journal of Organizational Behavior*, *32*(8), 1084-1105.

Van Velsor, E., Leslie, J. B., & Fleenor, J. W. (1997). *Choosing 360. A Guide to Evaluating Multi-Rater Feedback Instruments for Management Development*. Greensboro, NC: Center for Creative Leadership.

* Vassileva, I., Odlare, M., Wallin, F., & Dahlquist, E. (2012). The Impact of Consumers' Feedback Preferences on Domestic Electricity Consumption. *Applied Energy*, *93*, 575-582.

Verbeek, P. P. (2006). Persuasive Technology and Moral Responsibility Toward an Ethical Framework for Persuasive Technologies. *Persuasive*, *6*, 1-15.

Verplanken, B., & Faes, S. (1999). Good Intentions, Bad Habits, and Effects of Forming Implementation Intentions on Healthy Eating. *European Journal of Social Psychology*, *29*(5-6), 591-604.

Verplanken, B., & Wood, W. (2006). Interventions to break and create consumer habits. *Journal of Public Policy & Marketing*, *25*(1), 90-103.

* Wallenborn, G., Orsini, M., & Vanhaverbeke, J. (2011). Household appropriation of electricity monitors. *International Journal of Consumer Studies*, *35*(2), 146-152.

* Wang, S. C., & Chern, J. Y. (2013). Time-scheduled Delivery of Computer Health Animations: "Installing" Healthy Habits of Computer Use. *Health Informatics Journal*, *19*(2), 116-126.

Warnock, D., McGee-Lennon, M., & Brewster, S. (2011). The role of modality in notification performance. In: *Human-Computer Interaction–INTERACT 2011* (pp. 572-588). Springer Berlin Heidelberg.

* Willamowski, J. K., Hoppenot, Y., & Grasso, A. (2013). Promoting sustainable print behavior. In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1437-1442). ACM.

* Willis, R. M., Stewart, R. A., Panuwatwanich, K., Jones, S., & Kyriakides, A. (2010). Alarming visual display monitors affecting shower end use water and energy conservation in Australian residential households. *Resources, Conservation and Recycling*, *54*(12), 1117-1127.

Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.

Wood, W., & Neal, D. T. (2009). The habitual consumer. *Journal of Consumer Psychology, 19*, 579–592.

* Wood, G., & Newborough, M. (2003). Dynamic energy-consumption indicators for domestic appliances: environment, behavior and design. *Energy and Buildings*, *35*(8), 821-84

Yang, C. C., & Hsu, Y. L. (2010). A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, *10*(8), 7772-7788.

* Yu, Y., & Bhatti, S. N. (2014). The cost of virtue: reward as well as feedback are required to reduce user ICT power consumption. In *Proceedings of the 5th International Conference on Future Energy Systems* (pp. 157-169). ACM.

Zimmerman, B. J. (1998). Academic studing and the development of personal skill: A self-regulatory perspective. *Educational Psychologist*, *33*(2-3), 73-86.