Received: 31 January 2022 | Revised: 19 January 2024 | Accepted: 8 March 2024

DOI: 10.1111/isj.12520

SPECIAL ISSUE PAPER

WILEY

# Sustainable energy consumption behaviour with smart meters: The role of relative performance and evaluative standards

Charlotte Wendt[1] | Dominick Kosin[1] | Martin Adam[2] | Alexander Benlian[1]

[1]Institute of Information Systems and Electronic Services, Technical University of Darmstadt, Darmstadt, Germany

[2]Chair of Smart Services, Georg-August University of Göttingen, Göttingen, Germany

**Correspondence**
Charlotte Wendt, Institute of Information Systems and Electronic Services; Technical University of Darmstadt; Hochschulstr. 1, 64289 Darmstadt, Germany.
Email: charlotte.wendt@outlook.de

## Abstract

The growing adoption of smart meters enables the measurement of households' energy consumption, influenced not solely by building characteristics such as thermal insulation but also by residents' behavioural patterns, such as heating and ventilation practices. To motivate residents to adopt more sustainable behaviours, user interfaces on smartphones and laptops are increasingly using consumption data from households' smart meters to enable effective goal-setting. In contrast to previous research largely focusing on goal-setting in isolation, this study examines the role of specific social comparison-related design features that future research and practitioners can consider along with goal-setting to stimulate sustainable behaviours. Specifically, we look into the influence of residents' perception of their *relative performance* (i.e., whether their behaviour was better or worse than a reference group) on their ambition to act (i.e., targeted improvement goal) and their actual energy consumption behaviour. Moreover, we investigate the influence of a goal's *evaluative standard* (i.e., whether the goal refers to one's own or other's performance) on the relationship between relative performance, ambition to act, and energy consumption behaviour.

Drawing on social comparison theory, we conducted a framed field experiment with 152 households. We find that a goal's evaluative standard influences residents' awareness of their relative performance, affecting their ambition to act and, ultimately, their energy consumption behaviour. More specifically, we find that whereas other- (vs. self-) referencing goals encourage residents from worse-than-average performing households more strongly to improve their energy consumption behaviour, they discourage better-than-average ones. Overall, our study provides novel insights into the interplay between relative performance and evaluative standards as a means of fostering social comparison in smart meter-facilitated goal-setting, highlighting their crucial role in effectively supporting sustainable behaviours.

**KEYWORDS**

energy consumption, evaluative standard, goal-setting, green information systems, relative performance, smart meter

## 1 | INTRODUCTION

To transition to a more sustainable society, governmental policymakers increasingly call for information systems (IS) that support sustainable energy consumption, with smart meter technology leading the way (e.g., Cooper & Molla, 2017; Henkel & Kranz, 2018; Malhotra et al., 2013). Smart meters measure energy consumption and can communicate consumption data close to real-time through user interfaces (e.g., via smartphones or laptops), which can influence residents' energy consumption behaviours (e.g., heating, ventilation) (e.g., Sim et al., 2023; Tiefenbeck et al., 2018; Wunderlich et al., 2019). Residents in private households account for about 25% of the global energy consumption (Eurostat, 2021; IEA et al., 2020) and can benefit from this access to timely information to assess the impact of their behaviours. Driven by regulations, the current share of 35% of European households equipped with smart meters is expected to increase to 90% by 2030 (Alaton & Tounquet, 2020). As such, smart meters will allow for new levels of user engagement through user interfaces: As smart meters are further rolled out and advance in their capabilities, smart meter providers can design user interfaces in a way that encourages residents to conserve energy by setting themselves energy consumption goals (e.g., Abrahamse et al., 2007; Loock et al., 2013). Goal-setting in user interfaces is one of the most effective functionalities to encourage sustainable behaviours (e.g., Abrahamse et al., 2007; Harkin et al., 2016; Loock et al., 2013). In addition, smart meter providers can use social comparison elements such as performance indicators in the form of *relative (peer) performance* to inform residents whether their households' energy consumption is above or below a reference group (e.g., Allcott, 2011; Jetzek et al., 2014; Oracle Utilities, 2020). Smart meter user interfaces even allow incorporating such information about households' energy consumption and relative performance into the goal-setting process to encourage residents to set more ambitious goals and ultimately realise more sustainable energy consumption behaviours. Moreover, relative performance-based goal-setting can be leveraged to employ different *evaluative standards* of a goal (e.g., Elliot et al., 2011; Elliot & McGregor, 2001): *Self-referencing goals* measure success relative to one's previous performance (e.g., “I want to behave 10% *better than I did* last month”) whereas *other-referencing goals* measure success relative to the performance of peers (e.g., “I want to behave 10% *better than other comparable households do* on average”).

Due to the limited availability of timely data about comparable other households, either relative performance (i.e., social comparison) without goal-setting or self-referencing goals without relative peer performance used to be common means to stimulate energy consumption reduction. With the increasing proliferation of smart meters that enable access to timely data from various households and thus provide reliable data for reference groups, relative performance-based goal-setting as well as other-referencing goals that reinforce social comparison are now increasingly discussed (e.g., Gholami et al., 2016; Lossin et al., 2016; Wunderlich et al., 2019). Once this data is available more broadly, households can easily compare themselves with their peers based on their energy consumption (i.e., social comparison) and can be provided with self- and other-referencing goals. However, before smart meter providers engage with relative performance-based goal-setting, they need to better understand the effects of relative performance and evaluative standards as a means of fostering social comparison on residents' goal-setting and energy consumption behaviour. It is hereby particularly interesting to explore how the effectiveness of goal-setting may differ depending on the residents' relative performance (i.e., better vs. worse than the reference group) in combination with the employed evaluative standard (i.e., self- vs. other-referencing goal).

While investigating the effects of relative performance and evaluative standards on energy consumption behaviour is crucial for smart meter-enabled goal-setting and stimulating sustainable energy consumption, previous IS research has insufficiently explored these design features and their combinations. We specifically refer to Green IS research that has analysed design features in goal-setting and their potential to motivate residents to improve resource consumption behaviour (e.g., Gholami et al., 2016; Malhotra et al., 2013). From this existing literature, we have identified two important research gaps: First, previous research agrees that goal-setting generally has a positive impact on residents' energy consumption behaviour (e.g., Abrahamse et al., 2007; Loock et al., 2013) and initial theoretical conceptualizations indicate a positive effect of social comparison elements like relative performance in goal-setting on stimulating sustainable behaviours (Elliot et al., 2011; Lindenberg & Steg, 2007). However, information about one's relative performance and goal-setting have been treated separately in Green IS research, with their effects being investigated mainly in isolation. This is critical because the effect size for setting ambitious goals and related energy consumption behaviour may differ for residents with different performance levels. Specifically, the differentiation between over- and under-performing residents might be crucial given that prior research outside of goal-setting reveals that the awareness about one's relative performance (i.e., social comparison) encourages under- and over-performing individuals differently (e.g., Allcott, 2011; Loock et al., 2011; Schultz et al., 2007). Accordingly, current literature fails to capture important insights to better understand the effects of relative performance-based goal-setting on residents' energy consumption behaviour.

Second, previous research has mainly conceptualised goal-setting as a self-focused process in which residents' goals refer to their own performance without considering the actions of others (e.g., Abrahamse et al., 2007; Graham et al., 2011; Loock et al., 2013). As such, the goal's evaluative standard was mainly self-referencing, neglecting other-referencing goals (e.g., Elliot et al., 2011; Elliot & McGregor, 2001). However, a goal's evaluative standard may influence the relationship between relative performance and energy consumption behaviour. Residents' perception of their relative performance is likely to be reinforced when goals explicitly refer to their performance relative to peers, because such goals amplify the salience of social comparison. At the same time, the influence of an other-referencing goal on the effect of relative performance on residents' energy consumption behaviour is far from conclusive in the literature: On the one hand, other- (vs. self-) referencing goals may heighten peer performance awareness, spurring under-performing residents towards self-improvement and motivating over-performing residents to sustain their superior performance (e.g., Festinger, 1954; Lindenberg & Steg, 2007; Seidler et al., 2020; Wheeler, 1966). Conversely, they may discourage under-performers and lead over-performers to become over-confident and to prioritise comfort over pro-environmental behaviour (Festinger, 1954; Wills, 1981). As such, we require more nuanced insights into the influence of a goal's evaluative standard on the relationship between a household's relative performance and energy consumption behaviour. These insights are theoretically important as they help researchers better understand different goal-setting designs that affect sustainable behaviours. Without such knowledge, smart meter providers may assume that with access to more smart meter data, other-referencing goals are always better to encourage improved energy consumption behaviour, when in fact, self- and other-referencing goals both have their advantages, depending on the households' relative performances. Against this backdrop, we seek to answer the following research questions:

**RQ1.** What is the influence of smart meter information about households' relative (peer) performance (i.e., worse vs. better than average), provided at the time of goal-setting, on residents' energy consumption behaviour?

**RQ2.** What is the influence of a goal's evaluative standard (i.e., self- vs. other-referencing) on the relationship between households' relative performance and residents' energy consumption behaviour?

Drawing on social comparison theory, we conducted a framed field experiment – also referred to as lab-in-the-field experiment (e.g., Harrison & List, 2004; Karahanna et al., 2018) – with residents from 152 households. Participants reported their households' heating behaviour at three instances over 4 weeks while pursuing a self-set goal with an evaluative standard referring to either themselves or others. Our study reveals that residents exhibit more substantial improvements in energy conservation when they initially perform worse, rather than better, than their peers. This effect is mediated by residents' ambition to improve their energy consumption behaviour. Importantly, having an other- vs. a self-referencing goal leads to different results depending on residents' relative performance: Residents with subpar performance demonstrate greater improvement with other-referencing goals, whereas those initially outperforming their peers show decreased efforts with other-referencing goals as compared to self-referencing ones. In summary, our data indicate an estimated 11% to 24% enhancement in energy consumption, contingent upon households' initial performance relative to their peers.

With our study, we add to Green IS research (e.g., Gholami et al., 2016; Malhotra et al., 2013) and particularly to those investigating the use of Green IS to achieve sustainability outcomes (Kotlarsky et al., 2023) in two important ways: First, while previous research has looked at social comparison and goal-setting as two isolated functionalities for IS artefacts, we shed light on the influence of relative performance as a social comparison element for goal-setting on stimulating residents' energy consumption behaviour, but to different extents for over- versus under-performing residents. To that end, we also reveal that residents' ambition to act mediates the effect of relative performance in goal-setting on residents' energy consumption behaviour. This leads us to our second contribution: We enhance our understanding of how evaluative standards can be applied to smart meter-facilitated goal-setting to foster social comparison and to effectively motivate residents to achieve their targeted energy consumption behaviour. We shift the focus in the existing literature from understanding goal-setting primarily as a self-referral process to incorporating other-referencing goals as a complementary evaluative standard for goals. Specifically, we focus on contrasting the effects of self- and other-referencing goals on residents' perception of their relative performance (i.e., social comparison) at the time of goal-setting, which ultimately influences their energy consumption behaviour. Overall, we contribute to Green IS research by enhancing our understanding of the interplay between social comparison through relative performance and a goal's evaluative standard to stimulate sustainable behaviours. These insights also offer important and actionable implications for providers of smart meter-facilitated goal-setting, particularly regarding which evaluative standard is more effective in encouraging sustainable energy consumption behaviours when displaying relative performance.

## 2 | THEORETICAL BACKGROUND

In this section, we begin by reviewing prior research on the proliferation of smart meters as Green IS artefacts and position our study at the intersection of three literature streams: Green IS (with a focus on smart meters), goal-setting, and social comparison. Subsequently, we present social comparison theory as our theoretical lens through which we theorise about residents' perception of their relative performance at the time of goal-setting and how a goal's evaluative standard influences these perceptions and, ultimately, residents' energy consumption behaviour.

## 2.1 | Research on green IS and smart meters

Green IS research studies the design, implementation, and impact of technologies that facilitate sustainable behaviours to achieve pro-environmental objectives (e.g., Cooper & Molla, 2017; Henkel & Kranz, 2018; Kotlarsky et al., 2023; Watson et al., 2010). Our research is situated in the so far sparsely investigated category of impact-oriented Green IS research, that is the use of technologies and systems to achieve sustainability outcomes (e.g., Gholami et al., 2016; Malhotra et al., 2013). Green IS research in this area has begun to explore the role of smart meters in promoting sustainable behaviours by tracking and motivating individuals (e.g., Dalén & Krämer, 2017; Moors & Schäfer, 2023; Sim et al., 2023; Yang et al., 2021). Our study focuses on investigating social comparison in goal-setting, two concepts that have been predominantly studied in isolation.

Previous Green IS studies with a focus on smart meters investigated either (1) general drivers for the adoption of or resistance to smart meters (e.g., Chanson et al., 2019; Warkentin et al., 2017; Wunderlich et al., 2019); or (2) the intersection between smart meters and goal-setting to motivate sustainable behaviour with certain goal-related design features, such as the presence, absence, and variation of default goals, feedback on one's goal achievement, and monetary incentives to reach one's goal (e.g., Abrahamse et al., 2007; Graham et al., 2011; Loock et al., 2013); or (3) the interaction between smart meters and social comparison to stimulate sustainable behaviour, such as smart meter information on relative performance through normative feedback (e.g., Loock et al., 2011; Seidler et al., 2020). However, the intersection of all three fields has received only scant research attention, providing only conceptual and qualitative insights (e.g., Wendt & Benlian, 2022). Figure 1 visualises these intersections and provides exemplary publications for all relevant areas.

The intersection of smart meter-facilitated goal-setting with social comparison elements bears great potential for understanding the effects of peer performance on residents' goal-setting ambition and, ultimately, on their energy consumption behaviour. The resulting theoretical and practical implications are crucial, as the effects of social comparison in the context of smart meter-facilitated goal-setting may be contingent on social comparison elements. Without knowing the influence of social comparison elements on the relationship between residents' goal-setting and energy consumption behaviour, smart meter providers may extrapolate findings from either goal-setting or social comparison literature without considering insights from their combination. For example, smart meter providers may derive that one particular comparison element is always more effective when, in fact, the impact highly depends on the household's characteristics, such as its relative performance at the time of goal-setting. As a result, our research is novel and theoretically significant, as it deepens our understanding of the contingent effects of social comparison with smart meter-facilitated goal-setting.

## 2.2 | Social comparison theory and evaluative standards of goals

Social comparison literature examines the phenomenon of how an individual's behaviour is influenced by social norms and social influences, particularly through comparison with others or with oneself (e.g., Cialdini et al., 1991; Schultz et al., 2007). To that end, social comparison theory is grounded in the notion that individuals possess an inherent drive to evaluate themselves, for instance, in relation to their own performance (Festinger, 1954). In situations where objective means are not available or difficult to assess, individuals leverage the performance of others to evaluate their relative (peer) performance (i.e., whether one's own performance is better or worse than the performance of peers) and thus engage in social comparison (Festinger, 1954). As per the tenets of social comparison theory, the act of comparing oneself with others or oneself can yield two outcomes: First, when individuals compare

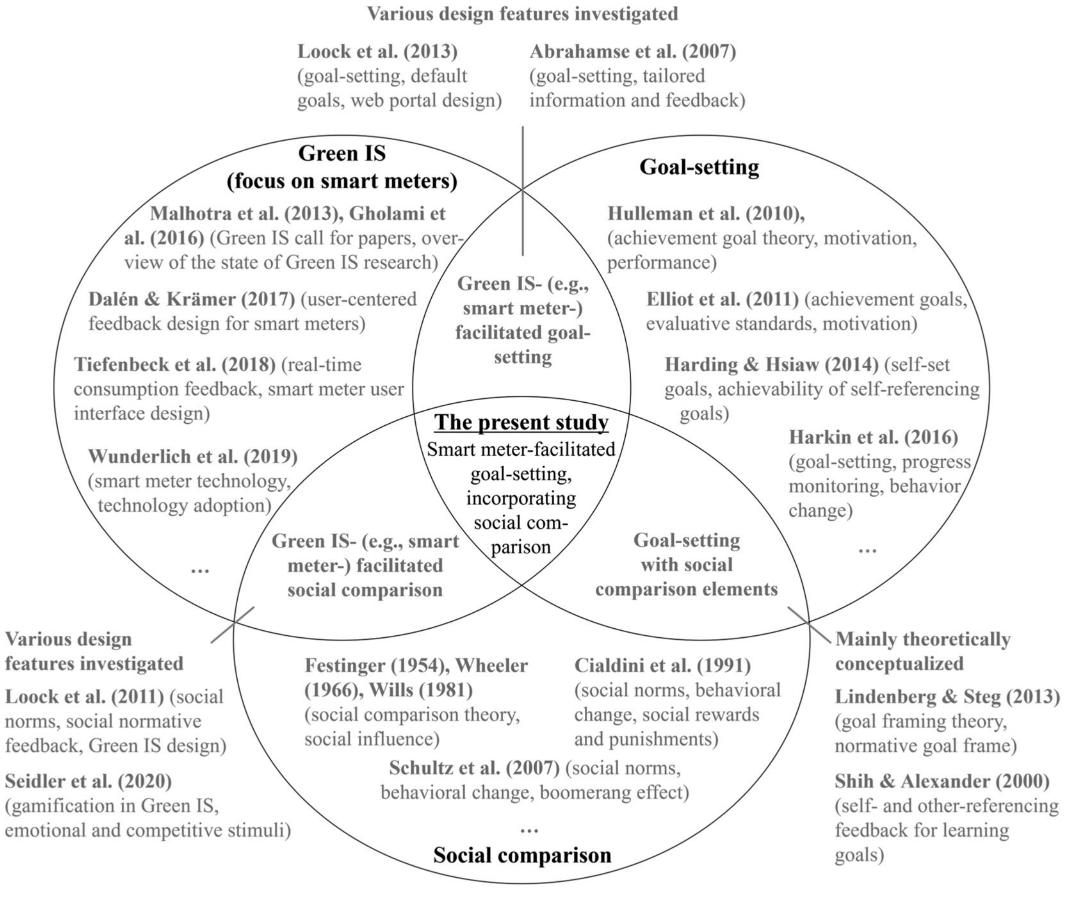


**FIGURE 1** Illustrative studies in the Green IS literature focusing on smart meters, goal-setting, and social comparison.

themselves to better-performing peers or a better past performance of themselves, they perceive their own current performance as comparatively inferior, leading them to engage in *upward comparison*. Second, when individuals compare themselves to worse-performing peers or a worse past performance of themselves, they perceive their own current performance as comparatively superior, giving rise to *downward comparison* (Festinger, 1954). Upward comparison typically raises the aspiration to achieve results similar to the reference value, thereby inducing self-improvement to match the norm (Festinger, 1954; Wheeler, 1966). This striving for a certain status also reflects the ambition of the individual to behave in a way that corresponds to their striving for the norm (Judge & Kammeyer-Mueller, 2012). In contrast, downward comparison results in pronounced self-esteem. Instead of striving for self-improvement, individuals tend to focus on failure avoidance (i.e., avoiding becoming worse than the norm) (Festinger, 1954; Wills, 1981).

Given its utility in explaining IS user behaviours, social comparison theory has been applied to various IS contexts within and beyond Green IS. For instance, previous research identified social comparison as an impactful design feature for persuasive systems (e.g., Corbett, 2013; Oinas-Kukkonen & Harjumaa, 2009) and studied how social comparison can be leveraged to nudge individuals towards sustainable behaviours (e.g., Staudt et al., 2021), specifically through smart meter-facilitated goal-setting (Loock et al., 2013; Wendt & Benlian, 2022).

To encourage social comparison, Green IS artefacts must offer comparative information, with the choice of a goal's evaluative standard being a promising design consideration. The evaluative standard of a goal distinguishes three different points of reference (i.e., self, others, and task) which can be used to evaluate if one is doing well or poorly while pursuing a goal (Elliot et al., 2011). Self-referencing goals use one's own development over time as the evaluative standard and define competence in terms of doing well or poorly relative to how one has performed in the past or will potentially perform in the future (e.g., one's performance this year vs. last year) (Elliot et al., 2011; Hulleman et al., 2010). In this regard, self-referencing goals leverage comparison elements by making the desired distinction from a past or future self more salient. Other-referencing goals use an interpersonal evaluative reference point and define competence in terms of doing well or poorly relative to comparable others (e.g., one's own performance vs. the average performance of others) (Elliot et al., 2011). Accordingly, other-referencing goals focus on demonstrating competence through comparison with peers and as such, by reinforcing social comparison (Hulleman et al., 2010; Nicholls, 1984). Importantly, individuals notice the performance of their peers not only by observing their peers' behaviours in direct interaction but also by having access to an IS that provides aggregated normative information, for instance displayed through user interfaces. Lastly, task-referencing goals use the absolute demands of a task as the evaluative standard and define competence in terms of performing well or poorly relative to what is required by the task itself (e.g., consuming less than a certain amount of a resource) (Elliot et al., 2011). Thus, they focus on mastering a particular task and attaining task-based competence (Elliot et al., 2011; Hulleman et al., 2010) and do not relate to social comparison, which is why they are not considered further in the context of our study.

Applied to our research context, we focus on smart meters as Green IS artefacts and how evaluative standards can be applied to smart meter-facilitated goal-setting to foster social comparison and to effectively motivate residents to achieve their targeted energy consumption behaviour. Specifically, we focus on contrasting the effects of self- and other-referencing goals on residents' perception of their relative performance (i.e., social comparison) at the time of goal-setting, which should ultimately influence their energy consumption behaviour. We do so for several reasons: Previous research on goal-setting in the context of energy consumption has mainly focused on self-referencing goals and related variables, such as the role of goal commitment and how realistic self-referencing goals are (e.g., Harding & Hsiaw, 2014), as well as feedback about self-referencing goal progress and achievement (e.g., Becker, 1978; Harkin et al., 2016; Lazaric & Toumi, 2022; van Houwelingen & van Raaij, 1989). In the same vein, Green IS studies, specifically those on smart meters, mainly analysed self-referencing goals, such as the role of self-referencing default goals and goal feedback (e.g., Abrahamse et al., 2007; Loock et al., 2013). However, previous Green IS research has largely overlooked crucial indications that highlight the potential effectiveness of the other-referencing evaluative standard as a design feature for promoting pro-environmental behaviours through social comparison. This evaluative standard will gain prominence with the advancements in smart meter technology and the increased data availability. It is particularly well-suited for inducing sustainable behaviours due to its explicit linkage of goal-setting to individuals' performance relative to peers (Lindenberg & Steg, 2007; Seidler et al., 2020).

According to recent research Wendt and Benlian (2022), there is evidence to suggest that conducting a comparative analysis of self- and other-referencing goals can offer valuable insights for enhancing the effectiveness of goal-setting facilitated by smart meters. However, it is worth noting that prior research in this area has primarily relied on qualitative approaches, lacking a more rigorous quantitative analysis. As such, with the quantitative research approach of this paper, we look more deeply into the difference between improving oneself over time (i.e., self-referencing goal) and improving compared to one's peers (i.e., other-referencing goal). Even though previous research on social comparison showed that peer performance can act as a magnet for both worse- and better-performing residents (e.g., Loock et al., 2011; Schultz et al., 2007), we know little about how this effect unfolds in goal-setting and which evaluative standard yields the best results for which group of residents. This research aims to gather empirical evidence to better understand how sustainable energy consumption behaviour can be effectively motivated through a combination of smart meter-facilitated goal-setting and social comparison elements.

# 3 | RESEARCH MODEL AND HYPOTHESES DEVELOPMENT

In the following, we utilise social comparison theory and goal-setting literature to formulate hypotheses that elucidate the influence of a goal's evaluative standard on the relationship between households' relative performance and their residents' energy consumption behaviour. We first explain the fundamental relationship between a household's relative performance at the time of goal-setting and its residents' subsequent energy consumption behaviour (H1). Next, we hypothesize how this relationship is mediated by residents' ambition to act (i.e., targeted goal to improve their energy consumption behaviour) (H2). As a consequence of these relationships, we can hypothesize how a goal's evaluative standard moderates the effects of relative performance on ambition to act (H3), and therefore how it can serve as a (personalisable) design feature for smart meter-facilitated goal-setting. Figure 2 depicts all hypotheses in a research model.

## 3.1 | The direct effect of relative performance on energy consumption behaviour

We first apply findings from social comparison literature to our context of smart meter-facilitated goal-setting to explain the fundamental relationship between households' relative performance and their residents' subsequent energy consumption behaviour. According to social comparison theory, individuals exhibit an inherent tendency to evaluate themselves regularly and leverage social comparison particularly in situations where objective means are not available or difficult to assess (Festinger, 1954). As such, social comparison strongly affects human behaviours and has repeatedly been demonstrated to play an important role, particularly in environmental contexts (e.g., Alberts et al., 2016; Allcott, 2011; Lossin et al., 2016), for example, when individuals compare their energy consumption to that of peers (i.e., relative performance) or try to reduce the discrepancy between themselves and the reference group of peers (Seidler et al., 2020; Staudt et al., 2021). To hypothesize the effect of such relative performance on energy consumption behaviour in the context of goal-setting, we differentiate between residents from households that perform worse and those who perform better than others at the time of goal-setting. We thereby draw on the average performance of peers (i.e., comparable other households) as a meaningful and salient reference point, which can be conveyed through the goal-setting functionality of smart meter user interfaces.

Residents from households whose relative performance is worse than average engage in upward comparison, which has been shown to raise the desire to achieve results (i.e., energy consumption levels) similar to those of one's peers (Loock et al., 2011; Schultz et al., 2007). Residents are encouraged to perform at the level established by the group norm. When they observe others performing better, this upward comparison will likely motivate them to improve themselves (Festinger, 1954). Such a comparative evaluation corresponds to the motivational effect associated with the process of setting oneself a goal (Locke & Latham, 2002). Accordingly, we argue that residents whose smart meter-facilitated user interface indicates worse-than-average performance at the time of goal-setting subsequently improve their energy consumption behaviour to reach levels close to or even better than average.

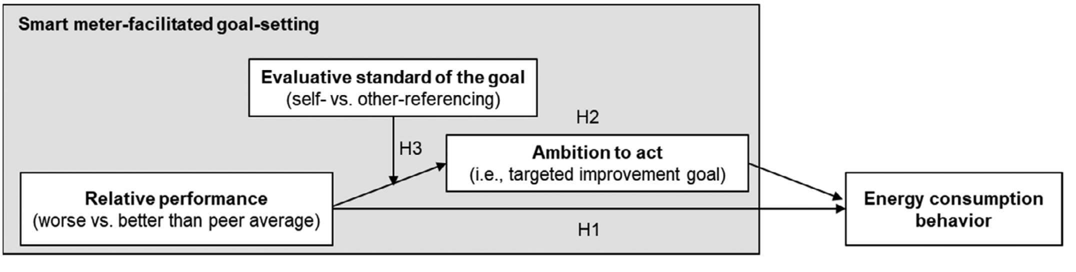


**FIGURE 2** Research model.

In contrast, residents from households that perform better than average experience a downward comparison. Consequently, these residents develop self-esteem and, in some cases, overconfidence, which can result in reduced motivation to further improve their behaviour once their superior performance becomes apparent through the user interface (Festinger, 1954; Wills, 1981). In this way, downward comparison can trigger the so-called boomerang effect, where the performance of residents from better-than-average households may not only stall but even deteriorate. For example, previous research showed that residents consuming less electricity than average increased their electricity consumption after being provided with feedback on their performance relative to peers (e.g., Loock et al., 2011; Schultz et al., 2007). Applying these insights to smart meter-facilitated goal-setting, where the process of goal-setting itself serves as a motivational lever, we contend that residents from households with below-average performance during goal-setting exhibit greater improvement in energy consumption behaviour compared to those from households with above-average performance, who still improve but with less effort. As such, we hypothesize:

**H1.** Residents from households whose smart meter-facilitated user interface indicate a worse- (vs. better-) than-average performance at the time of goal-setting subsequently improve their energy consumption behaviour to a higher degree.

## 3.2 | The mediating role of ambition to act

A deeper understanding of the underlying mechanism that accounts for the previously theorised effect is vital to unravel the implications of relative performance on residents' energy consumption behaviour in smart meter-facilitated goal-setting. Generally, individuals who freely set a goal strive to adapt their behaviours to an extent that allows them to meet their chosen goal (Harkin et al., 2016). In that sense, an individual's self-set goal reflects their ambition to act (e.g., to improve their energy consumption behaviour), referring to an individual's "persistent and generalised striving for success, attainment, and accomplishment" (Judge & Kammeyer-Mueller, 2012, p. 759). More challenging goals should thus reflect greater ambition to act (e.g., Harding & Hsiaw, 2014; Hirschi & Spurk, 2021; Loock et al., 2013). Moreover, the ability to compare one's own performance against that of peers and, therefore, to engage in social comparison is one of the main levers driving individuals' ambition to act (Festinger, 1954; Wills, 1981).

Applied to the context of smart meter-facilitated goal-setting, residents' ambition to act is stimulated when they are given visibility into their energy consumption behaviour relative to peers (i.e., social comparison) while setting a goal through the user interface. By observing their relative performance, residents are encouraged to adopt more sustainable behaviours and ultimately decrease their energy consumption (Loock et al., 2011), which is reflected in their self-set goal. In that sense, the goal set by a resident is a tangible and objective indicator of their ambition to act (e.g., Harding & Hsiaw, 2014; Loock et al., 2013). However, the level of ambition to act (i.e., the targeted goal) is likely different for residents whose smart meters inform them that their relative performance is worse- compared to better-than-average: We expect the level of ambition of residents who engage in upward comparison (i.e., from worse-than-average households) to be high because they strive for self-improvement and accordingly set a goal to improve their consumption behaviour considerably. In contrast, we expect residents who engage in downward comparison (i.e., from better-than-average households) to exhibit comparably lower levels of ambition to engage in energy-conserving behaviours and, thus, to set a goal to maintain or even worsen their consumption behaviour. Therefore, we hypothesize:

**H2.** Residents' ambition to act mediates the effect of relative performance on energy consumption behaviour such that residents whose smart meter-facilitated user interface indicates a worse- (vs. better-) than-average performance at the time of goal-setting exhibit higher ambition to act and improve their energy consumption behaviour to a higher degree.

### 3.3 | The moderating role of a goal's evaluative standard

Considering the notable impact of goal-setting on guiding individuals' attention and effort towards goal-relevant activities (Locke & Latham, 2002), we aim to hypothesize on the role of evaluative standards in fostering social comparison in goal-setting. Specifically, we explore how evaluative standards influence residents' assessment of cognitively processed information, shaping their orientation towards self or others, and ultimately motivating them to take action (Elliot & McGregor, 2001; Lindenberg & Steg, 2007).

Self-referencing goals cause residents to compare their performance with standards of their own likening (Elliot et al., 2011). To that end, the propensity to improve one's own performance is primarily based on intrinsic motivation drivers such as personal interest (e.g., protecting the environment), a strive to master the required actions (e.g., airing intermittently), or the intent to do better when compared to the past (Hulleman et al., 2010). In this sense, self-referencing goals are likely to lead residents to focus on their self-improvement potential. Smart meter information on relative performance is likely to be used merely to determine a reasonable level for the self-improvement goal. In other words, the performance of peers is used as an anchor.

In contrast, other-referencing goals aim to demonstrate one's competence relative to others and thus are based on a normative component (Hulleman et al., 2010). While smart meter information on one's relative performance encourages social comparison rather implicitly, this normative component in an other-referencing goal makes social comparison more explicit and salient (Hulleman et al., 2010). Thereby, normative concerns (i.e., the fear of performing worse than others) tend to evoke higher levels of ambition to act when setting a goal (Hansson et al., 1983) and serve as a lever to stimulate pro-environmental behaviours (Lindenberg & Steg, 2007). As such, we argue that other-referencing goals intensify residents' awareness of and attention to peer performance and emphasise social comparison. Instead of peer performance merely serving as an anchor for a reasonable goal, it now directly influences the achievement of the targeted goal. As a result, we expect that information about peers' energy consumption is processed more consciously by residents who set themselves other-referencing goals compared to those who set self-referencing goals.

Accordingly, we contend that a goal's evaluative standard significantly impacts how relative performance and as such social comparison affects energy consumption behaviour through ambition to act. In other words, the evaluative standard of a goal moderates the influence of relative performance on residents' ambition to act. Specifically, for residents performing worse than average, the heightened focus on peers likely accentuates normative concerns. As a result, the ambition to act of residents with an other-referencing (vs. self-referencing) goal increases (i.e., setting a higher reduction goal), leading to a more substantial improvement in energy consumption behaviour. Similarly, we expect that an other-referencing (vs. self-referencing) goal causes residents that perform better than average to become more aware of their currently superior standing. Accordingly, these residents develop fewer normative concerns and should be less ambitious (i.e., setting a laxer goal), ultimately affecting their energy consumption behaviour. Combined with the mediation hypothesized in H2, we thus propose the following moderated mediation hypothesis:

> **H3.** The effect of relative performance on energy consumption behaviour via ambition to act is moderated such that a goal's evaluative standard influences the effect of relative performance on ambition to act.

## 4 | METHOD

We tested our hypotheses by conducting a 2 × 2 (relative performance: worse vs. better than average; evaluative standard of the goal: self- vs. other-referencing) framed field experiment (e.g., Fink, 2022; Harrison & List, 2004; Karahanna et al., 2018). Framed field experiments use a sample of real subjects in actual settings where participants make real-life decisions related to the study. Therefore, framed field experiments "leverage the subjects' real-world settings for the context to increase external validity and realism" (Karahanna et al., 2018, p. iv) and are thus particularly suitable to investigate not only behavioural intentions but also actual behaviours.

## 4.1 | Experimental context

We chose to investigate smart meter-facilitated goal-setting for motivating sustainable energy consumption in the highly relevant context of (space) heating: With a share of 64% of households' energy consumption in the EU (Eurostat, 2020), heating of households constitutes an exceptional potential for energy savings. Apart from costly investments into building infrastructure, it is assumed that residents can actively reduce heating energy (HE) consumption by about 20% simply through more energy-conserving behaviours, such as airing intermittently and closing shades at night (e.g., Alberts et al., 2016; Pfnür & Müller, 2021; Vassileva et al., 2013).

We chose participants from Germany as saving energy is of high political and societal priority in Germany, particularly as heating in Germany is largely powered by oil and gas (e.g., Wunderlich et al., 2019). Moreover, residents from German households are affected by the EU's smart meter directives and will comprehensively be supplied with smart meters and regular consumption information in the near future (Alaton & Tounquet, 2020). Besides, given that awareness of HE consumption behaviour is particularly relevant in countries with seasonal weather conditions (e.g., Catalina et al., 2008), we opted for Germany as a northern hemisphere country that is subject to seasonal variations. At the same time, German residents are not particularly aware of short-term energy price fluctuations, as heating bills are typically sent out only once per year.

Moreover, we decided to investigate self-reported heating *behaviour* instead of measuring absolute HE consumption data for two reasons: First, absolute HE consumption figures usually include fractions of total building consumption for households in shared buildings or rental apartments with collective heating systems, which individual households do not have full control over and access to (Eurostat, 2013). Second, given our focus on HE consumption reduction related to behavioural changes, reported heating behaviours are unaffected by potentially strong differences in consumption due to variations among households in structural measures and appliances (e.g., more or less insulation, different kinds of heaters) as well as an apartment's location (e.g., south vs. north side, detached vs. adjacent to other apartments) (e.g., Abrahamse et al., 2007; Vassileva et al., 2013).

## 4.2 | Experimental design and manipulations

Similar to previous studies on energy consumption behaviours (e.g., Sütterlin et al., 2011; Wemyss et al., 2019), we designed a custom-made online user interface where participants were asked to briefly report on their household's current heating behaviours. To evaluate households' inherent relative performance and to compare the effect of different evaluative standards of goals, these reported behaviours were used to build a single individual score, which we refer to as a household's heating intensity (e.g., Calì et al., 2016; Paone & Bacher, 2018). To increase participants' comprehensibility of the collected data, we described this heating intensity to participants as a normalised measure of their heating behaviour ranging from 1 to 100. Operationalising HE consumption in that way offered the advantage of providing participants with a tangible and comparable visualisation of their heating behaviours. Additionally, it allowed us to ensure that participants would not attribute any differences between their and other households' heating behaviour to structural factors (e.g., building insulation, old heating systems) that we cannot control for. We computed the heating intensity of each household to derive an overall average that would serve participants to assess their relative performance (RP). We decided to simplify the visualisation of this RP in that we employed only two categories for RP. Worse-than-average performing participants were shown to perform 20% worse than average, whereas better-than-average performing participants were shown to perform 20% better than average. This simplification allowed us to ensure that participants were aware that there was a non-negligible difference between

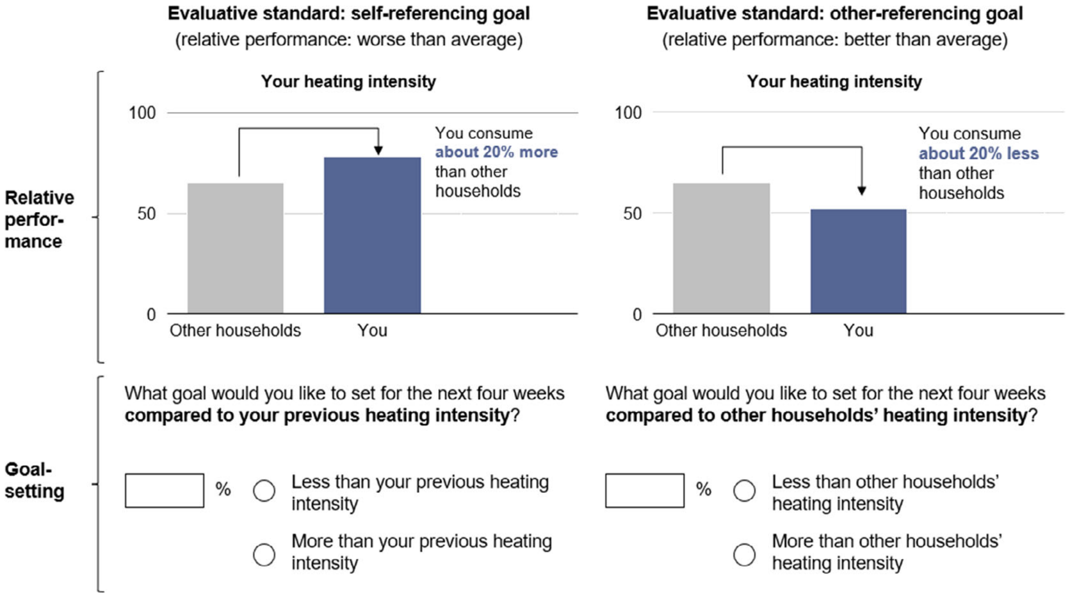


**FIGURE 3** Visualisation of smart meter information on relative performance along with manipulations for the goal's evaluative standard.

the average and their own heating intensity, while still maintaining comparability among all participants being better than average and all participants being worse than average. We opted for the number of 20% as it is assumed that this share can realistically be saved simply through more energy-conserving behaviours (e.g., airing intermittently, closing shades at night) (e.g., Alberts et al., 2016; Pfnür & Müller, 2021; Vassileva et al., 2013). With our simplified approach, we comparably encouraged all participants to engage in upward or downward comparison.

The experimental design also provided a goal-setting functionality to allow participants to set a self-selected goal for the upcoming 4 weeks that referred to either their own (i.e., self-referencing) or comparable others' heating intensity (i.e., other-referencing) as an evaluative standard (ES). The experimental duration of 4 weeks captures a stable and generalizable estimate of social lives (Reis & Wheeler, 1991, p. 287). Furthermore, participants received feedback on their goal attainment at the end of the study period. To assess goal attainment for participants who set an other-referencing goal, we needed a reference group to help us understand how the heating behaviour of the alleged group of others changed during the same period. On that account, we collected data from a control group who reported their heating behaviour, but did not receive any information on their RP and were not given the possibility to set a goal. Figure 3 depicts our visualisation of smart meter information on participants' RP (i.e., worse vs. better than average) and our operationalization of goal-setting with two different ES (i.e., self- vs. other-referencing goal).

## 4.3 | Experimental procedure

Figure 4 shows the experimental procedure for participants of all conditions: (1) First, we offered participants a clearer context for our study and introduced them to the setting by illustrating the importance of HE consumption behaviour with regard to overall energy consumption in households. This introductory information ensured that all participants were able to put themselves in the real situation of private heating during the experiment and understood the relevance of the topic, which is in line with current industry practices to raise awareness of sustainable energy consumption behaviour (e.g., Oracle Utilities, 2020). (2) Second, participants reported their own household's heating behaviour of the past 2 weeks, for instance, with respect to airing habits and usage of heaters. Participants were instructed to consider the collective behaviour of their entire household, emphasising the importance of reflecting upon the actions of the household as a unified entity, rather than focusing solely on their individual behaviours. (3) Next, according to their self-reported heating behaviour, participants were categorised into the group of households with either worse- or better-than-average RP and were shown the respective visualisation (see Figure 3). (4) Lastly, we randomly assigned participants to a condition that offered a goal-setting functionality with either self- or other-referencing ES where participants set their goal and confirmed their choice. (5) We asked participants to briefly report on their household's current heating behaviour after 2 weeks before inviting participants to the final reporting round after 4 weeks. (6) Based on the goal participants set themselves and their subsequent modifications in heating behaviour during the 4-week period, we proceeded to provide participants feedback on their goal attainment. In addition, a control group went through steps (1), (2), and (5) but received neither information about their relative performance nor a goal-setting functionality.

## 4.4 | Variables measured

We measured our dependent variable, heating intensity reduction, by calculating the difference between heating intensity in the initial round and 4 weeks later (Abrahamse et al., 2007). Our independent variable, relative performance, was assigned to participants according to their initial heating intensity based on their self-reported heating behaviour (i.e., worse- vs. better-than-average RP). For our mediating variable (i.e., ambition to act), we measured participants' chosen goals as an individual's self-set goal reflects their ambition to improve their energy consumption behaviour (e.g., Harding & Hsiaw, 2014; Hirschi & Spurk, 2021; Loock et al., 2013). To allow for comparability of goals between all treatment groups and to conduct our analyses, we converted other-referencing goals into equivalent self-referencing goals. For example, a household initially behaving 20% worse than others with a goal of eventually behaving 10% better than others had a goal that was equivalent to a self-referencing goal of 30% improvement. This self-referent goal value represents a household's ambition to act in that higher (vs. lower) goals indicate higher ambition to improve the current HE consumption behaviour.

To assess participants' heating intensity, we measured each household's self-reported heating behaviour using several seven-point Likert-type scale items (see Table A1 in the Appendix A). We then consolidated these items for further analyses into our heating intensity variable which we presented in the smart meter use interface as a scale ranging from 1 to 100 for participants' easier comprehension (e.g., Abrahamse et al., 2007; Sütterlin et al., 2011; Wemyss et al., 2019). To derive participants' heating intensity, we first asked participants about their possession and use of household appliances (Abrahamse et al., 2007). Specifically, we asked them to indicate which type of heater they owned (e.g., radiator with manual vs. programmable thermostat) and what temperature they chose during day- and night-time in the living- and bedroom (Guerra-Santin & Itard, 2010). Moreover, participants were asked to indicate how often their household performed different activities related to heating and airing habits (Calì et al., 2016; Sütterlin et al., 2011), for example "keep doors closed between rooms with different temperatures" as well as about their household's general awareness with regard to heating-related activities such as "closing shutters or blinds at night" (Paone & Bacher, 2018). We offered participants the possibility of choosing "not applicable" as a response option for questions that were specific to certain appliances (e.g., floor heating) (Sütterlin et al., 2011).

Following prior literature (e.g., Sütterlin et al., 2011; Wemyss et al., 2019), we computed each household's heating intensity by calculating the mean of the items for self-reported behaviour. Averaging participants' responses over all items also allowed for minimising potential measurement errors introduced by the discrete levels of our Likert-type scales. We then used this heating intensity to infer the overall average performance as well as households' RP and to assess their goal attainment at the end of our experiment. In addition, we measured demographics (i.e., participants' age and gender), household size (i.e., number of household members), intermediate performance (i.e., heating intensity after 2 weeks) as well as competitiveness (i.e., how strongly participants generally compare their performance) (Tiefenbeck et al., 2018) and social desirability bias (Paulhus, 1991) as control variables that we considered to be influential in our experimental setting. Table A1 in the Appendix A lists all employed items.

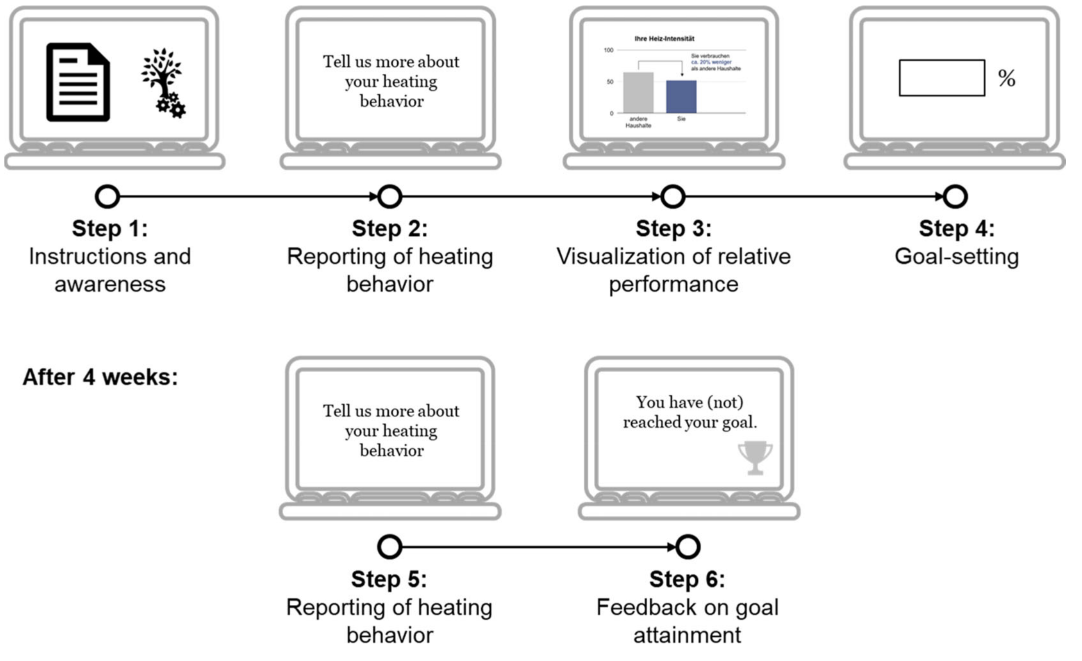


**FIGURE 4** Experimental procedure.

Additionally, we integrated three attention checks to notice whether participants read each item carefully and understood the meaning of the visualisations. Moreover, we asked about extended absences from home (e.g., Tiefenbeck et al., 2018) as well as how accurately participants perceived the employed survey items to capture and represent their actual heating behaviour.

## 5 | DATA ANALYSIS AND RESULTS

### 5.1 | Sample

We recruited 192 participants from the crowdsourcing platform Prolific which is specifically designed for behavioural research and experiments. Multiple studies corroborate that the data of Prolific respondents exhibits high reliability and is of comparable or even higher quality than responses from the commonly used online research platform Amazon Mechanical Turk (Palan & Schitter, 2017; Peer et al., 2017). We restricted access to our study to participants located in Germany with an approval rating above the suggested threshold of 95% to ensure high data quality (Goodman & Paolacci, 2017). Out of those 192 participants, we removed 40 who either did not answer all questions, failed our attention checks, or stated that they had been absent from their homes for more than 3 days during our experiment. As such, we arrived at a final data set with 152 valid entries. We collected the data from January to February 2021 to capture a timeframe in which heating is particularly relevant in Germany. Table A2 (see Appendix A) summarises the descriptive statistics of the participating households in the control group (i.e., no RP information, no goal-setting) and the two treatment groups (i.e., self- vs. other-referencing ES of goal), segmented by the corresponding RP (i.e., worse vs. better than average).

The results of several one-way analyses of variance for the control variables indicate good comparability and balance across not only our two treatment conditions (i.e., self- vs. other-referencing ES of the goal), but also across all sub-groups covering both worse- and better-than-average performance, as there were no significant differences ($p > 0.1$) in terms of participants' gender, age, household size, competitiveness or their stated social desirability bias. Therefore, we find support for the successful randomisation of the assignment process to our experimental conditions. In addition, we verified that participants whose heating behaviour was categorised as better than average did not exhibit any significant difference ($p > 0.1$) in the above-mentioned control variables compared to those participants categorised as worse than average – thereby ensuring comparability of both performance categories. Lastly, our respondents strongly agreed that the employed survey items accurately captured and represented their actual heating behaviour (M = 6.20, SD = 1.26).

## 5.2 | Influence of social desirability bias

As our dependent variable heating intensity behaviour was self-reported, we took particular care to address any influence resulting from participants describing their heating behaviour more favourably than might be the case – often referred to as social desirability bias. Accordingly, we followed established and contemporary best practices in research (e.g., Larson, 2019; Nederhof, 1985; Paulhus, 1991) and specifically in IS research (e.g., Turel et al., 2011) to minimise and account for potential social desirability bias: First, we designed our experiment to reduce the likelihood of social desirability bias by assuring participants that their results would not be assessed individually but only in aggregation with all other participants, by allowing them to set any goal they liked – purposefully integrating an option for a goal to behave even *worse* than they did before – and by not incentivising participants whatsoever in any change of heating behaviour. Therefore, participants were neither formally nor informally pressured or constrained in their goal-setting and attainment. Second, after collecting our data, we used participants' responses to the social desirability bias items and compared social desirability bias across all four groups (i.e., ES: self- vs. other-referencing, RP: worse vs. better than average), finding no significant differences ($p > 0.1$) and thus indicating that social desirability bias did not distort the outcomes across all four groups. Third, we included the social desirability bias measure as a control variable in our regressions, accounting for any potential effects in the analyses. Fourth and lastly, we calculated Spearman's correlation between the social desirability bias measure and our dependent variable, heating intensity behaviour. Since Spearman's correlation was insignificant ($p > 0.1$), the results suggest that social desirability bias has no association with participants' responses on their heating intensity behaviour. Overall, we conclude that social desirability bias – as far as it occurred at all – does not meaningfully affect the implications of our study.

## 5.3 | Contrast analysis of heating intensity reduction

To gain an initial understanding of the effects of households' RP and goals' ES on households' heating behaviour via their heating intensity reduction, we performed a contrast analysis based on independent samples *t*-tests. Performing better (vs. worse) than average at the time of goal-setting decreases heating intensity reduction by 53.7% (8.8% vs. 19.0%; $p < 0.01$), thereby indicating a significant influence of RP on heating intensity reduction. To obtain a more differentiated view, we consider the goal's ES: As visualised in Figure 5, among those participants pursuing a self-referencing goal, we do not find any significant difference in heating intensity reduction between worse- and better-than-average performing participants (13.5% vs. 11.1%; $p > 0.1$). However, for those participants following an other-referencing goal, the difference between worse- and better-than-average performing participants reaches 73.6% (24.1% vs. 6.4%; $p < 0.001$). As such, we find an initial indication that the degree to which RP impacts heating intensity reduction is subject to the ES of the goal pursued.

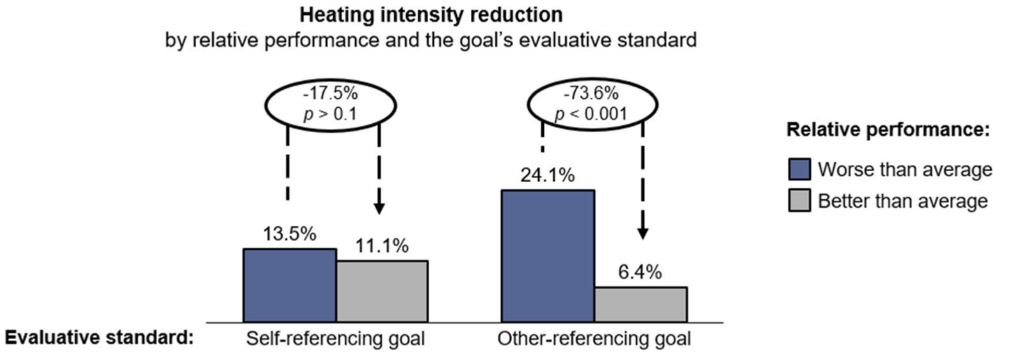


**FIGURE 5** Effect of a goal's evaluative standard on heating intensity reduction, subject to households' relative performance.

## 5.4 | Hypotheses testing

We first conducted a hierarchical linear, ordinary least squares regression to test our hypotheses, as reported in Table 1. Focusing on our dependent variable heating intensity reduction, we designed four stages. In stage 1, we included our control variables gender, age, household size, social desirability, competitiveness, and intermediate performance as well as our independent variables RP (worse-than-average RP $= 0$, better-than-average RP $= 1$) and ES (self-referencing ES $= 0$, other-referencing ES $= 1$). In stage 2, we added the term for the interaction between RP and ES. In Stage 3, we analysed the mediating influence of ambition to act. We furthermore analysed a fourth stage in which we included ambition to act as our mediator and ES as our moderator. Similarly, we set up the previously described stage 1 and 2 also for our mediator ambition to act (i.e., the targeted improvement goal).

The results support H1, indicating that households with a better-than-average RP achieve a smaller reduction in heating intensity compared to those with worse-than-average RP (stage 1: $\beta = -10.46$, $p < 0.01$). Moreover, our findings reveal that ambition to act significantly predicts heating intensity reduction (Stage 3: $\beta = 0.67$, $p < 0.001$), providing the first indication for mediation stated in H2 (Baron & Kenny, 1986). Similarly, we find an initial indication for our moderated mediation mentioned in H3 in that the interaction of RP and ES significantly influences the ambition to act (stage 2: $\beta = -9.68$, $p < 0.05$).

Subsequently, we verified our mediation hypothesis (H2) and our moderated mediation hypothesis (H3). Starting with ambition to act as a hypothesized mediator carrying over the effect of RP on heating intensity reduction, we conducted a bootstrap analysis with 5000 bootstrap samples and 95% confidence intervals using PROCESS (Hayes, 2022, Model 4), which further supports the mediating effect of ambition to act (see Table 2). As the regression results in Table 1 indicate that including ambition to act as a mediator leads to no significant influence of the direct effect of RP on heating intensity anymore, we find support for an indirect-only mediation (Zhao et al., 2010). Next, we analysed the moderating effect of ES on the indirect effect of RP on heating intensity reduction by conducting a bootstrap analysis with 5000 bootstrap samples and 95% confidence intervals using PROCESS (Hayes, 2022, Model 8). The results in Table 2 support that ES indeed interacts with RP in the form of a moderated mediation. As such, we find evidence supporting both the mediation hypothesized in H2 and the moderated mediation hypothesized in H3. Hence, the effect of RP in which worse- (vs. better-) than-average households show higher ambition to reduce their heating intensity is reinforced when the goal's ES switches from self-referencing to other-referencing.

Figure 6 depicts the interaction plots for both ambition to act and heating intensity reduction.

**TABLE 1** Hierarchical linear regression on ambition to act and heating intensity reduction.

| | Ambition to act (i.e., targeted reduction goal) | | Heating intensity reduction | | | |
|---|---|---|---|---|---|---|
| | Stage 1 | Stage 2 | Stage 1 | Stage 2 | Stage 3 | Stage 4 |
| Manipulations | | | | | | |
| RP | −11.16*** (2.43) | −6.39† (3.37) | −10.46** (3.00) | −2.57 (4.09) | −3.86 (3.05) | 1.18 (3.87) |
| ES | 5.37* (2.42) | 10.44** (3.48) | 5.31† (3.13) | 13.91** (4.37) | 5.05† (2.85) | 11.17** (4.08) |
| Interaction | | | | | | |
| RP × ES | – | −9.68* (4.83) | – | −16.12** (5.88) | – | −11.42* (5.53) |
| Mediation | | | | | | |
| Ambition to act | – | – | – | – | 0.67*** (0.14) | 0.61*** (0.14) |
| Control variables | | | | | | |
| Household size | −1.15 (0.91) | −0.88 (0.91) | 0.62 (1.12) | 1.05 (1.10) | 1.17 (1.03) | 1.43 (1.02) |
| Gender | −1.48 (2.60) | −1.68 (2.56) | 1.28 (3.21) | 0.97 (1.84) | 2.78 (2.93) | 2.43 (2.89) |
| Age | 0.13 (0.20) | 0.20 (0.20) | −0.52* (0.25) | −0.40 (0.24) | −0.48 (0.23) | −0.39 (0.23) |
| Social desirability bias | −5.38** (1.53) | −5.21 (1.51) | 0.15 (1.93) | 0.36 (1.88) | 2.26 (1.81) | 2.23 (1.78) |
| Competitiveness | −1.10 (1.52) | −1.35 (1.51) | 1.36 (1.88) | 0.93 (1.84) | 1.61 (1.71) | 1.28 (1.69) |
| Intermediate performance | – | – | 0.27*** (0.07) | 0.29*** (0.07) | 0.51*** (0.08) | 0.50*** (0.08) |
| $R^2$ | 0.28 | 0.31 | 0.27 | 0.32 | 0.40 | 0.42 |

*Note*: $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; standard errors in parentheses.
Abbreviations: ES, evaluative standard.; RP, relative performance.

## 6 | DISCUSSION

Excessive energy consumption is one of the most pressing sustainability problems in the 21st century, with households accounting for more than 25% of total energy consumption (Eurostat, 2021; IEA et al., 2020). Our research aimed to explore the impact of smart meter information about households' relative performance in energy consumption on their residents' energy consumption behaviour. By doing so, we sought to enhance the effectiveness of motivating residents to improve their energy footprint. Moreover, we set out to examine how evaluative standards – as a smart meter-facilitated design feature for goal-setting – shape the influence of relative performance on residents' energy consumption behaviour. Specifically, we explored how the inclusion of peer performance in the goal-setting process affects this relationship. Our research reveals that the configuration of a goal's evaluative standard (i.e., whether one's goal refers to oneself or peers) significantly shapes residents' awareness of their relative performance and thus has a positive impact on their energy consumption behaviour. More specifically, we find that a goal's evaluative standard influences the mediated effect of a household's relative performance on residents' energy consumption behaviour via ambition to act: Residents from worse-than-average households are more ambitious and thus improve their energy consumption behaviour more strongly with an other-referencing goal (estimated improvement by 24.1%) than with a self-referencing goal (estimated improvement by 13.5%). In contrast, residents from better-than-average households decrease their efforts to improve their energy consumption behaviour and show less improvement with an other-referencing goal (estimated improvement by 6.4%) than with a self-referencing goal (estimated improvement by 11.1%).

**TABLE 2** Bootstrap analyses for the (moderated) mediation relative performance → ambition to act → heating intensity reduction.

| | Moderator | Effect | $Boot_{SE}$ | $Boot_{LLCI}$ | $Boot_{ULCI}$ |
|---|---|---|---|---|---|
| Mediation | – | Direct: 3.82 | 3.08 | −2.29 | 9.92 |
| | | Indirect: 6.64 | 1.84 | 3.38 | 10.59 |
| Moderated mediation | ES: self-referencing | Direct: −1.18 | 3.87 | −8.86 | 6.49 |
| | | Indirect: 3.75 | 1.88 | 0.54 | 7.87 |
| | ES: other-referencing | Direct: 10.24 | 4.31 | 1.70 | 18.77 |
| | | Indirect: 8.45 | 2.63 | 4.02 | 14.22 |
| | Index of moderated mediation | | $Boot_{SE}$ | $Boot_{LLCI}$ | $Boot_{ULCI}$ |
| | 4.70 | | 2.61 | 0.08 | 10.32 |

*Note*: 5000 bootstrap samples, 95% confidence intervals using PROCESS (Hayes, 2022).
Abbreviation: ES: evaluative standard.

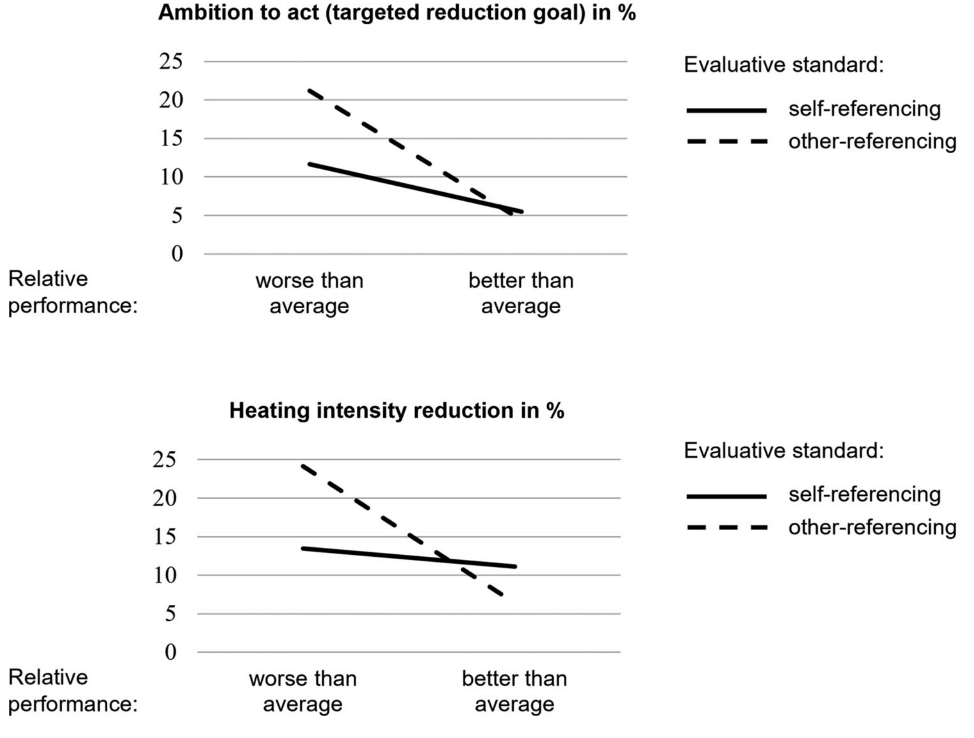


**FIGURE 6** Interaction between relative performance and a goal's evaluative standard.

## 6.1 | Contributions to green IS research

Our study offers two important contributions to Green IS research. It particularly adds to the Green IS stream on the use of technologies and systems to achieve sustainability outcomes (cf. Kotlarsky et al., 2023), which has repeatedly emphasised the need for empirical exploration of how Green IS artefacts effectively promote sustainable behaviours (e.g., Gholami et al., 2016; Malhotra et al., 2013).

First, we add to the body of knowledge on Green IS by showing how incorporating households' relative performance as a social comparison element into the goal-setting process affects residents' energy consumption behaviour. Previous Green IS literature outside the context of goal-setting reveals that smart meter information on one's relative performance stimulates only residents from under-performing households to improve their energy consumption behaviour while it encourages residents from over-performing households even to worsen their energy consumption behaviour (e.g., Loock et al., 2011; Schultz et al., 2007). At the same time, theoretical conceptualizations about integrating normative information (e.g., relative performance) into the goal-setting process assume a predominantly positive effect on residents' energy consumption behaviour but overlook to distinguish residents based on their relative performance (Elliot et al., 2011; Lindenberg & Steg, 2007). Our findings confirm these theoretical conceptionalisations and show that providing information about a household's relative performance at the time of goal-setting influences the effectiveness of smart meter-facilitated goal-setting and thus, residents' energy consumption behaviour. Beyond this, our results reveal that the overall positive effect of goal-setting is not the same for all residents – as implicitly assumed by previous goal-setting research (e.g., Abrahamse et al., 2007; Loock et al., 2013). Instead, the positive effect changes in its effect size depending on households' relative performance at the time of goal-setting: Residents from under-performing households improve their energy consumption behaviour to a higher degree than those from over-performing households. As such, we reveal that relative performance as a social comparison element determines the effectiveness of goal-setting and ultimately the achievement of sustainable outcomes. Thereby, we also add to the social comparison literature which revealed that peer performance acts as a magnet for both worse- and better performing individuals, meaning that individuals who already perform better than average decrease their effort and start to perform worse once they become aware of their superior performance (e.g., Allcott, 2011; Loock et al., 2011; Schultz et al., 2007). With our study we increase our understanding of how social comparisons can cause a positive effect (i.e., improved behaviour) on both, worse- and better-performers when being combined with goal-setting.

Moreover, we uncover an important underlying mechanism that explains the effectiveness of incorporating relative performance into the process of goal-setting. Prior Green IS literature on goal-setting mainly followed a black-box-approach by narrowing its focus to quantifying the extent to which goal-setting design features lead to an objective change, such as absolute energy consumption reduction (e.g., Abrahamse et al., 2007; Loock et al., 2013). Our study goes beyond previous research by revealing that residents' ambition to act mediates the effect of one's relative performance in goal-setting on energy consumption behaviour, with residents from worse-than-average households setting more ambitious goals and improving their energy consumption behaviour to a higher degree than residents from better-than-average households. Furthermore, by using residents' self-set goals as a measure of their ambition to act, we not only underscore the substantial relevance of goal-setting as an effective smart meter functionality. Our research also enhances our comprehension of how considering social comparison through one's relative performance influences residents' behaviour when setting and pursuing a goal. As such, our insights are valuable to mitigate the risk of implementing designs that cause behavioural changes in unintended directions, as has been observed for individuals who lack the possibility to set a goal and who start to neglect pro-environmental behaviours once they become aware that their relative performance is better than that of peers (e.g., Loock et al., 2011; Schultz et al., 2007).

Second, our research enhances the field of Green IS by deepening our understanding of a goal's evaluative standard as a customizable design element in goal-setting that fosters social comparison. The evaluative standard of a goal significantly influences the salience of social comparison elements and thus, the extent to which households' relative performance affects their residents' ambition to act, ultimately impacting their energy consumption behaviour. Previous Green IS literature investigated various design features that potentially influence the efficacy of goal-setting, such as default goals, goal feedback, and goal incentives (e.g., Abrahamse et al., 2007; Graham et al., 2011; Loock et al., 2013; Yim, 2011). In doing so, prior literature implicitly conceptualised goal-setting as a process in which individuals focus solely on their own performance, with goals generally being self-referencing (e.g., Abrahamse et al., 2007; Loock et al., 2013) and social comparison either not being present (e.g., goal incentives) or limited to comparing oneself to one's own previous or future performance (e.g., default goal, goal feedback). In contrast, other-referencing goals have largely been neglected. We introduce a goal's evaluative standard (i.e., self- vs. other-referencing goals) as a design feature that – in the case of other-referencing goals – explicitly links individuals' goals to the performance of peers, thereby reinforcing relative performance as a social comparison element in goal-setting. Although other-referencing goals have been discussed conceptually, their empirical investigation remains unexplored (Elliot et al., 2011; Lindenberg & Steg, 2007). As such, our study deviates from and expands upon prior research in Green IS by challenging the assumption that goals are relatively static and self-centred targets. Instead, we explicitly juxtapose two different reference points for goal-setting (i.e., one's own vs. comparable others' performance) and investigate their effectiveness for motivating and achieving sustainable behaviours. To that end, we examine how linking the reference point of a goal to the performance of residents' peers influences residents' goal choice and their energy consumption behaviour. This perspective sheds light on how social comparison-related design features for goal-setting affect the way residents engage in pro-environmental behaviours. This is important as it unlocks new and powerful ways to motivate residents in setting and pursuing ambitious goals through smart meter-facilitated goal-setting. More specifically, our results demonstrate that other-referencing goals intensify residents' consideration of the performance of their peers (i.e., social comparison), meaning that the highest improvement in energy consumption behaviour is achieved when user interfaces offer residents from under-performing households a goal with an other-referencing evaluative standard, and residents from over-performing households a goal with a self-referencing evaluative standard. As such, our study provides valuable insights for Green IS artefacts such as smart meter-enabled goal-setting seeking to inspire and foster more sustainable behaviours.

## 6.2 | Practical implications for IS and sustainable energy consumption behaviour

Our study offers actionable insights for smart meter providers and policymakers seeking to encourage sustainable energy practices among residents, by delving into the interplay between relative performance (i.e., social comparison), evaluative standards, and energy consumption behaviour. The pivotal role of evaluative standards as a means of fostering social comparison in goal-setting becomes evident when residents assess their household's performance against themselves or peers while setting energy consumption goals. To leverage this, designers and providers of smart meter user interfaces are encouraged to personalise goal-setting through evaluative standards, tailoring them to households' relative performance. For households performing worse than average, presenting other-referencing goals proves most effective, while households performing better than average benefit from self-referencing goals. With the increasing accessibility of smart meter data, households' relative performance can be continuously assessed by smart meter providers, facilitating direct customization of evaluative standards in smart meter user interfaces based on households' unique conditions.

Furthermore, smart meter providers can enhance the effectiveness of their devices' user interfaces by offering a more targeted reference point for other-referencing goals than the simple average. For example, for households whose relative performance is superior to that of the average of comparable households, designers and providers of smart meter user interfaces could present a goal referring to the top quartile of comparable households (i.e., "I want to behave 10% better than the top quartile of my peers"). This nuanced personalization approach ensures that residents from high-performing households are challenged appropriately, fostering ambitious yet realistic goal-setting. By strategically incorporating different evaluative standards of goals in smart meter user interfaces, providers and designers can effectively utilise the power of evaluative standards to foster social comparison and to steer residents towards sustainable energy consumption. In practice, following our recommendations can lead to a substantial 11% to 24% improvement in energy consumption, contributing to reduced emissions and costs. As smart meters become more widespread globally, our findings gain more practical significance and utility, offering valuable insights for ongoing and future efforts in sustainability initiatives.

### 6.3 | Limitations and directions for future research

While our study offers several contributions and implications, it is nevertheless subject to limitations. First, our results are based on a framed field experiment, whereby some aspects (i.e., relative performance) were derived from participants' reported real-life behaviour whereas other aspects (i.e., evaluative standard of the goal) were randomly assigned to participants. Even though this type of experiment allows for gathering data from participants' self-reported behaviours, participants are aware of the experimental setting. To account for potential social desirability bias, we followed established and contemporary procedures to minimise and control any potential confounding influences. As our tests revealed, social desirability poses no noteworthy implications for our research findings. However, to improve the generalizability and confirm the transferability and robustness of our findings, we call for future research to conduct field studies (e.g., field or quasi-field experiments) where participants are less aware of the research design. Conceivably, such a field study could additionally collect objective consumption data on energy consumption (e.g., by tapping into timely and even real-time smart meter data) to eliminate potential limitations of self-reported measurements and to capture energy consumption changes on a continuous scale without the limitations inherent to self-reported and discrete measurements. Similarly, future studies can investigate how self-reported consumption behaviour and smart meter-tracked absolute consumption data diverge and may potentially influence goal-setting. Moreover, scholars may gather additional data on residents' behaviours by examining their engagement with smart meters, such as frequency of usage, as well as exploring the potential impact of these behaviours on other resource consumption patterns, like the frequency and duration of showering. In this vein, our choice of ambition to act as a mediator was motivated to offer a parsimonious research model while not overstraining the paper's complexity, specifically given our moderation mediation analyses. Therefore, future studies may investigate additional mediators (e.g., goal difficulty) to provide a richer explanation of the underlying mechanisms.

Second, our experiment comprised a period of 4 weeks. To account for seasonal fluctuations beyond volatile day-to-day changes driven by varying outside temperatures, we encourage future research to investigate how the effects of a goal's evaluative standard unfold over more extended periods (e.g., several heating seasons). Such insights would help to improve our understanding of whether and how residents uphold their efforts to conserve energy when they relate to others that experience the same external fluctuations (in the case of other-referencing goals) rather than to themselves, where they may constantly struggle to select a reasonable goal (in the case of self-referencing goals). In the same vein, it would be interesting to investigate how smart meter providers can switch the same household's goal design when the household adjusts its energy consumption behaviours over the long term. Specifically, once a household moves from a worse- to a better-than-average performance, the goals' evaluative standard could switch from other-referencing to self-referencing to ensure a higher ambition to act and a further improved energy consumption behaviour. In this regard, future investigations could delve into alternative reference points for a goal's evaluative standard, such as contrasting past-month versus past-week behaviours or comparing households based on averages versus the top 10% performers. Additionally, there may be greater scope for enhancing sustainable energy consumption in households performing below average, warranting focused studies on optimising interventions for this specific target group.

Third, our sample included German participants. Even though Germans are particularly affected by smart meter directives and experience seasonal weather conditions that make the investigated design feature (i.e., evaluative standards of goals) instrumental, other countries and cultures, especially beyond Europe, may behave differently. As such, we encourage future research also to investigate other cultural contexts to complement and corroborate the findings from our experiment. In doing so, scholars can advance our understanding of further implications of different evaluative standards of goals and account for moderating factors (e.g., preferences of different household members). Moreover, future scholarly investigations could employ more comprehensive face-to-face interviews to yield even richer insights.

Lastly, to our knowledge, our study is one of the first to shed light on how different evaluative standards of goals can be leveraged to curb residents' energy consumption behaviour. Nevertheless, Green IS research on goal-setting could gain further insights from a more nuanced distinction within and between self- and other-referencing goals. Future research can explore the influence of perceived group proximity on goal-setting and goal attainment among closely related residents. Specifically, it would be valuable to investigate how residents perceive the group of others in relation to themselves. Understanding the impact of this perceived group proximity on the process of setting and achieving goals can provide valuable insights into individual and collective behaviours. Essentially, goals' evaluative standards represent an impactful opportunity waiting to be applied and tested by (Green) IS researchers in many contexts that aim to influence behaviours well beyond heating behaviour.

In conclusion, the study of smart meter-facilitated goal-setting holds significant promise for addressing the pressing global need to motivate residents to adopt sustainable resource consumption behaviours and transition towards a more environmentally conscious society. We hope our study is one step towards making a more sustainable future a reality.

## ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

Data available on request from the authors.

## ORCID

*Charlotte Wendt* https://orcid.org/0000-0003-2271-5253
*Dominick Kosin* https://orcid.org/0000-0002-4605-5395
*Martin Adam* https://orcid.org/0000-0001-9369-7203
*Alexander Benlian* https://orcid.org/0000-0002-7294-3097

## REFERENCES

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2007). The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. *Journal of Environmental Psychology*, *27*(4), 265–276. https://doi.org/10.1016/j.jenvp.2007.08.002

Alaton, C., & Tounquet, F. (2020). *Benchmarking smart metering deployment in the eu-28: Final report*. European Commission Publications Office. Retrieved from https://data.europa.eu/doi/10.2833/492070

Alberts, G., Gurguc, Z., Koutroumpis, P., Martin, R., Muûls, M., & Napp, T. (2016). Competition and norms: A self-defeating combination? *Energy Policy*, *96*, 504–523. https://doi.org/10.1016/j.enpol.2016.06.001

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Becker, L. J. (1978). Joint effect of feedback and goal setting on performance: A field study of residential energy conservation. *Journal of Applied Psychology*, *63*(4), 428–433.

Calì, D., Osterhage, T., Streblow, R., & Müller, D. (2016). Energy performance gap in refurbished german dwellings: Lesson learned from a field test. *Energy and Buildings*, *127*, 1146–1158. https://doi.org/10.1016/j.enbuild.2016.05.020

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage Publications.

Catalina, T., Virgone, J., & Blanco, E. (2008). Development and validation of regression models to predict monthly heating demand for residential buildings. *Energy and Buildings*, *40*(10), 1825–1832. https://doi.org/10.1016/j.enbuild.2008.04.001

Chanson, M., Bogner, A., Bilgeri, D., Fleisch, E., & Wortmann, F. (2019). Blockchain for the IOT: Privacy-preserving protection of sensor data. *Journal of the Association for Information Systems*, *20*(9), 1274–1309. https://doi.org/10.17705/1jais.00567

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevalutation of the role of norms in human behavior. *Advances in Experimental Social Psychology*, *24*, 201–234. https://doi.org/10.1016/S0065-2601(08)60330-5

Cooper, V., & Molla, A. (2017). Information systems absorptive capacity for environmentally driven is-enabled transformation. *Information Systems Journal*, *27*(4), 379–425. https://doi.org/10.1111/isj.12109

Corbett, J. (2013). Designing and using carbon management systems to promote ecologically responsible behaviors. *Journal of the Association for Information Systems*, *14*(7), 339–378. https://doi.org/10.17705/1jais.00338

Dalén, A., & Krämer, J. (2017). Towards a user-centered feedback design for smart meter interfaces to support efficient energy-use choices – A design science approach. *Business & Information Systems Engineering*, *59*(5), 361–373. https://doi.org/10.1007/s12599-017-0489-x

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3 × 2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632–648. https://doi.org/10.1037/a0023952

Eurostat. (2013). *Manual for statistics on energy consumption in households*. Publications Office of the European Union.

Eurostat. (2020). *Energy consumption in households*. Retrieved from 04-10-2020: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households

Eurostat. (2021). *Sdg 7 – Affordable and clean energy: Ensure access to affordable, reliable, sustainable and modern energy for all*. Retrieved from 03-12-2021: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=SDG_7_-_Affordable_and_clean_energy#Energy_consumption

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Fink, L. (2022). Why and how online experiments can benefit information systems research. *Journal of the Association for Information Systems*, *23*(6), 1333–1346.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Gholami, R., Watson, R., Hasan, H., Molla, A., & Bjorn-Andersen, N. (2016). Information systems solutions for environmental sustainability: How can we do more? *Journal of the Association for Information Systems*, *17*(8), 521–536. https://doi.org/10.17705/1jais.00435

Goodman, J., & Paolacci, G. (2017). Crowdsourcing consumer research. *Journal of Consumer Research*, *44*(1), 196–210. https://doi.org/10.1093/jcr/ucx047

Graham, J., Koo, M., & Wilson, T. D. (2011). Conserving energy by inducing people to drive less. *Journal of Applied Social Psychology*, *4*(1), 106–118. https://doi.org/10.1111/j.1559-1816.2010.00704.x

Guerra-Santin, O., & Itard, L. (2010). Occupants' behaviour: Determinants and effects on residential heating consumption. *Building Research and Information*, *38*(3), 318–338. https://doi.org/10.1080/09613211003661074

Hair, J. F., Jr., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (pls-sem)* (2nd ed.). Sage Publications.

Hansson, R. O., Hogan, R., Johnson, J. A., & Schroeder, D. (1983). Disentangling type a behavior: The roles of ambition, insensitivity, and anxiety. *Journal of Research in Personality*, *17*(2), 186–197. https://doi.org/10.1016/0092-6566(83)90030-2

Harding, M., & Hsiaw, A. (2014). Goal setting and energy conservation. *Journal of Economic Behavior & Organization*, *107*, 209–227. https://doi.org/10.1016/j.jebo.2014.04.012

Harkin, B., Webb, T. L., Chang, B. P., Prestwich, A., Conner, M., Kellar, I., Benn, Y., & Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *142*(2), 198–229. https://doi.org/10.1037/bul0000025

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009–1055. https://doi.org/10.1257/0022051043004577

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis* (3rd ed.). Guilford Press.

Henkel, C., & Kranz, J. (2018). *Pro-environmental behavior and green information systems research-review, synthesis and directions for future research*. Paper presented at the ICIS 2018 Proceedings, San Francisco, CA, USA.

Hirschi, A., & Spurk, D. (2021). Striving for success: Towards a refined understanding and measurement of ambition. *Journal of Vocational Behavior*, *127*(103), 577. https://doi.org/10.1016/j.jvb.2021.103577

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422–429. https://doi.org/10.1037/a0018947

IEA, IRENA, UNSD, World Bank, & WHO. (2020). *Tracking sdg 7: The energy progress report 2020*. World Bank.

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014). Generating sustainable value from open data in a sharing society. In B. Bergvall-Kåreborn & P. A. Nielsen (Eds.), *Creating value for all through it. Tdit 2014* (pp. 62–82). Springer.

Judge, T. A., & Kammeyer-Mueller, J. D. (2012). On the value of aiming high: The causes and consequences of ambition. *Journal of Applied Psychology*, *97*(4), 758–775. https://doi.org/10.1037/a0028084

Karahanna, E., Benbasat, I., Bapna, R., & Rai, A. (2018). Editor's comments: Opportunities and challenges for different types of online experiments. *MIS Quarterly*, *42*(4), iii–x.
Kotlarsky, J., Oshri, I., & Sekulic, N. (2023). Digital sustainability in information systems research: Conceptual foundations and future directions. *Journal of the Association for Information Systems*, *24*(4), 936–952.
Larson, R. B. (2019). Controlling social desirability bias. *International Journal of Market Research*, *61*(5), 534–547.
Lazaric, N., & Toumi, M. (2022). Reducing consumption of electricity: A field experiment in Monaco with boosts and goal setting. *Ecological Economics*, *191*(107), 231.
Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of Social Issues*, *63*(1), 117–137. https://doi.org/10.1111/j.1540-4560.2007.00499.x
Locke, E., & Latham, G. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037/0003-066x.57.9.705
Loock, C.-M., Staake, T., & Landwehr, J. (2011). *Green is design and energy conservation: An empirical investigation of social normative feedback*. Paper presented at the ICIS 2011 Proceedings, Shanghai, China.
Loock, C.-M., Staake, T., & Thiesse, F. (2013). Motivating energy-efficient behavior with green is: An investigation of goal setting and the role of defaults. *MIS Quarterly*, *37*(4), 1313–1332. https://doi.org/10.25300/misq/2013/37.4.15
Lossin, F., Loder, A., & Staake, T. (2016). Energy informatics for behavioral change: Increasing the participation rate in an it-based energy conservation campaign using social norms and incentives. *Computer Science – Research and Development*, *31*(3), 149–155. https://doi.org/10.1007/s00450-014-0295-3
Malhotra, A., Melville, N. P., & Watson, R. T. (2013). Spurring impactful research on information systems and environmental sustainability. *MIS Quarterly*, *37*(4), 1265–1274. https://doi.org/10.25300/misq/2013/37:4.3
Moors, M., & Schäfer, R. (2023). Sustainability in private households-investigating acceptance of smart energy apps.
Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280.
Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91*(3), 328–346. https://doi.org/10.1037/0033-295X.91.3.328
Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. In J. C. Nunnally & I. H. Bernstein (Eds.), *Psychometric theory* (pp. 248–292). McGraw-Hill.
Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems*, *24*(1), 485–500. https://doi.org/10.17705/1cais.02428
Oracle Utilities. (2020). *Welcome to the future of home energy reports, today*. Retrieved from https://www.oracle.com/a/ocom/docs/industries/utilities/utilities-opower-home-energy-reports.pdf
Palan, S., & Schitter, C. (2017). Prolific.Ac – A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004
Paone, A., & Bacher, J.-P. (2018). The impact of building occupant behavior on energy efficiency and methods to influence it: A review of the state of the art. *Energies*, *11*(4), 953–971.
Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). Academic Press.
Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006
Pfnür, A., & Müller, N. (2021). *Die rolle des mieters im wohnungswirtschaftlichen klimaschutz*. Retrieved from https://noventic.com/media/pages/magazin/mieter-studie/0c9cd4cfed-1632473126/kurzfassung_noventic-studie_rolle-mieter-im-klimaschutz_210923.pdf
Reis, H. T., & Wheeler, L. (1991). Studying social interaction with the rochester interaction record. *Advances in Experimental Social Psychology*, *24*(3), 269–318. https://doi.org/10.1016/S0065-2601(08)60332-9
Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x
Seidler, A.-R., Henkel, C., Fiedler, M., Kranz, J., Ixmeier, A., & Strunk, K. S. (2020). *Promoting eco-sustainable behavior with gamification: An experimental study on the alignment of competing goals*. Paper presented at the ICIS 2020 Proceedings, A Virtual AIS Conference.
Sim, J., Lee, J., & Cho, D. (2023). On the effectiveness of smart metering technology adoption: Evidence from the national rollout in the United Kingdom. *Journal of the Association for Information Systems*, *24*(2), 555–591.
Staudt, P., Greif-Winzrith, A., & Nieken, P. (2021). *Increasing contributions to sustainable projects through digital nudges amplifying social comparison*. Paper presented at the ICIS 2021 Proceedings, Austin, Texas, USA.
Sütterlin, B., Brunner, T. A., & Siegrist, M. (2011). Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics. *Energy Policy*, *39*(12), 8137–8152. https://doi.org/10.1016/j.enpol.2011.10.008

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., & Staake, T. (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management Science*, *64*(3), 1458–1476. https://doi.org/10.1287/mnsc.2016.2646

Turel, O., Serenko, A., & Giles, P. (2011). Integrating technology addiction and use: An empirical investigation of online auction users. *MIS Quarterly*, *35*(4), 1043–1061. https://doi.org/10.2307/41409972

van Houwelingen, J. H., & van Raaij, W. F. (1989). The effect of goal-setting and daily electronic feedback on in-home energy use. *Journal of Consumer Research*, *16*(1), 98–105. https://doi.org/10.1086/209197

Vassileva, I., Dahlquist, E., Wallin, F., & Campillo, J. (2013). Energy consumption feedback devices' impact evaluation on domestic energy use. *Applied Energy*, *106*, 314–320. https://doi.org/10.1016/j.apenergy.2013.01.059

Warkentin, M., Goel, S., & Menard, P. (2017). Shared benefits and information privacy: What determines smart meter technology adoption? *Journal of the Association for Information Systems*, *18*(11), 3.

Watson, R., Boudreau, M.-C., & Chen, A. J. (2010). Information systems and environmentally sustainable development: Energy informatics and new directions for the is community. *MIS Quarterly*, *34*(1), 23–38. https://doi.org/10.2307/20721413

Wemyss, D., Cellina, F., Lobsiger-Kägi, E., de Luca, V., & Castri, R. (2019). Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland. *Energy Research & Social Science*, *47*, 16–27. https://doi.org/10.1016/j.erss.2018.08.018

Wendt, C., & Benlian, A. (2022). *Motivating sustainable resource consumption through the design of goal setting in smart meter user interfaces*. Paper presented at the European Conference on Information Systems.

Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, *1*(suppl.1), 27–31.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, *90*(2), 245–271. https://doi.org/10.1037/0033-2909.90.2.245

Wunderlich, P., Veit, D. J., & Sarker, S. (2019). Adoption of sustainable technologies: A mixed-methods study of german households. *MIS Quarterly*, *43*(2), 673–691. https://doi.org/10.25300/misq/2019/12112

Yang, Y.-C., Ying, H., Jin, Y., Cheng, H. K., & Liang, T.-P. (2021). Special issue editorial: Information systems research in the age of smart services. *Journal of the Association for Information Systems*, *22*(3), 10.

Yim, D. (2011). *Tale of two green communities: Energy informatics and social competition on energy conservation behavior*. Paper presented at the AMCIS 2011 Proceedings, Detroit, USA.

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering baron and kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206. https://doi.org/10.1086/651257

## AUTHOR BIOGRAPHIES

**Charlotte Wendt** holds a PhD in information systems management from the Technical University of Darmstadt, Germany. Her research interests include user decision-making under the influence of real-time information and goal-setting as well as digital transformation of organisations. Her work has been published in international journals such as the European Journal of Information Systems, Decision Support Systems, and Information Systems Frontiers, as well as in conferences such as the International Conference on Information Systems (ICIS) or the European Conference on Information Systems (ECIS).

**Dominick Kosin** holds a PhD in information systems management from the Technical University of Darmstadt, Germany. His research interests include user empowerment, web customization, and digital platforms. His work has been published in international journals such as the Information Systems Journal, the European Journal of Information Systems, and Electronic Markets as well as in conferences such as the International Conference on Information Systems (ICIS) or the European Conference on Information Systems (ECIS).

**Martin Adam** is a Chaired Professor of Information Systems, especially Smart Services, at University of Goettingen, Germany. His research interests include user-AI collaborations as well as the digital transformation of work and people. His work has been published in international journals, such as Information Systems Journal, Information Systems Research, Journal of the AIS, European Journal of Information Systems, Decision Support Systems, Communications of the Association for Information Systems, Business & Information Systems Engineering, and Electronic Markets. He serves as Associate Editor at Information Systems Journal and Electronic Markets.

**Alexander Benlian** is a Chaired Professor of Information Systems, especially Electronic Services, at Technical University of Darmstadt, Germany. His research focuses on the digital transformation of work and people, IT entrepreneurship, algorithmic management in organisational and platform environments, and human-computer interaction. His research has been published in MIS Quarterly, Information Systems Research, Information Systems Journal, Journal of Management Information Systems, Journal of the AIS, Journal of Information Technology, Entrepreneurship Theory & Practice, Journal of Business Venturing, Journal of Service Research, and MIS Quarterly Executive, among others. He serves as Senior Editor at European Journal of Information Systems, as Department Editor at Business & Information Systems Engineering, and as Editorial Board Member at Information Systems Research and MISQ Executive.

**How to cite this article:** Wendt, C., Kosin, D., Adam, M., & Benlian, A. (2024). Sustainable energy consumption behaviour with smart meters: The role of relative performance and evaluative standards. *Information Systems Journal*, *34*(6), 1995–2023. https://doi.org/10.1111/isj.12520

## APPENDIX A

**TABLE A1** Constructs and measurement items.

| Construct | Items |
|---|---|
| Self-reported heating behaviour | Heating appliances and use (Guerra-Santin & Itard, 2010)<br>What thermostats do you mainly use in your living room / in your bedroom?<br>What temperature are the thermostats (/what level are the radiators) currently set to in your *living room* during the day?<br>What temperature are the thermostats (/what level are the radiators) currently set to in your *bedroom* during the night?<br>Heating habit (Sütterlin et al., 2011)<br>What share of your home do you currently heat? [no room at all … all rooms]<br>How often do you currently keep doors closed between rooms of different temperatures? [very rarely … very often]<br>How long per day do you use additional (mostly electric) heating elements such as heating mats, electric blankets or fan heaters? [(almost) not at all … more than 3 h; not applicable]<br>Airing habit (Calì et al., 2016)<br>How long per day do you currently leave your windows tilted? [(almost) not at all … more than 3 h; not applicable]<br>How long per day do you currently leave your windows (wide) open? [(almost) not at all … more than 3 h; not applicable]<br>General awareness (Paone & Bacher, 2018)<br>How often do you currently close shutters or blinds at night? [(almost) not at all … very often; not applicable]<br>How often do you currently make sure your radiators are vented? [(almost) not at all … very often; not applicable]<br>How much attention do you pay to your furniture not shielding your radiators? [(almost) not at all … very often; not applicable] |
| Social desirability bias (Paulhus, 1991) | My first impression of people usually turns out to be right.<br>I am often not confident of my judgement. (r)<br>I always know why I like things.<br>I have once received too much change from a salesperson without telling him/her. (r)<br>I never tell lies to others.<br>There have been occasions when I have taken advantage of someone. (r)<br>(first three items for self-deception enhancement, followed by three items for impression management, all measured from strongly disagree to strongly agree) |
| Competitiveness (Tiefenbeck et al., 2018) | In general, I often evaluate my performance in comparison to the performance of others.<br>In general, I often compare my performance to my own past performance. |

*Note*: (r): reversed item.

### A.1. | MEASUREMENT MODEL ASSESSMENT OF OUR CONTROL VARIABLE SOCIAL DESIRABILITY BIAS

We assessed the psychometric properties of our measurement models by examining convergent validity for our construct social desirability bias, which we verified using three criteria (Fornell & Larcker, 1981): First, the item loadings were significant ($p < 0.001$) and exceeded the recommended level of 0.70 (Carmines & Zeller, 1979). Second, with a value of 0.94 for Cronbach's alpha and 0.94 for composite reliability, both measures of internal consistency were above the recommended level of 0.70 (Nunnally & Bernstein, 1994). Third, with a value of 0.72, the average variance extracted exceeded the relevant threshold of 0.50 (Hair. et al., 2016). Hence, our construct for social desirability bias met the specifications for convergent validity and was appropriate to be employed as a control variable to account for respondents' desire to avoid embarrassment and project a favourable image to others.

The results of several one-way analyses of variance for the control variables indicate good comparability and balance across not only our two treatment conditions (i.e., self- vs. other-referencing ES of the goal), but also across all sub-groups covering both worse- and better-than-average performance, as there were no significant differences ($p > 0.1$) in terms of participants' gender, age, household size, competitiveness or their stated social desirability bias. Therefore, we find support for the successful randomisation of the assignment process to our experimental conditions, which is also in line with prior research on the use of smart meters, finding no significant influence of demographic variables on individuals' aspiration to consume energy in a sustainable fashion (Dalén & Krämer, 2017).

**TABLE A2** Descriptive statistics.

| Variable \ Group | Control ($N = 36$) | | Self-referencing ES: Worse-than-average RP ($N = 27$) | | Self-referencing ES: Better-than-average RP ($N = 31$) | | Other-referencing ES: Worse-than-average RP ($N = 29$) | | Other-referencing ES: Better-than-average RP ($N = 29$) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD | M | SD |
| Household size | 2.75 | 1.76 | 2.70 | 1.59 | 2.39 | 1.26 | 2.38 | 1.15 | 2.76 | 1.41 |
| Gender (Female) | 25% | – | 33% | – | 32% | – | 41% | – | 34% | – |
| Age | 28.08 | 9.21 | 28.81 | 6.78 | 28.61 | 7.81 | 25.45 | 5.14 | 29.34 | 6.24 |
| Social desirability bias | 4.63 | 0.86 | 4.66 | 0.81 | 4.58 | 0.84 | 4.71 | 0.70 | 4.89 | 0.83 |
| Competitiveness | 3.43 | 0.54 | 3.04 | 0.79 | 3.31 | 0.91 | 3.04 | 0.90 | 3.34 | 0.90 |
| Chosen reduction goal | – | – | 11.67% | 13.01% | 5.48% | 11.72% | 21.21% | 15.62% | 4.38% | 12.08% |
| Heating intensity reduction | 4.54% | 19.01% | 13.47% | 19.91% | 11.11% | 15.60% | 24.15% | 14.70% | 6.37% | 16.76% |

Abbreviations: ES, evaluative standard; M, mean; RP, relative performance; SD: standard deviation.