behavior, energy & climate change

BECC 2021

Eco-feedback

CJ Meinrenken
Columbia University

cmeinrenken@ei.columbi.edu

IN LVMINE TVO VIDEBIMVS LVMEN

Abstract

As shown in past studies, providing household residents with information about their electricity usage can prompt reductions. However, some questions remain regarding how such feedback needs to be designed to prompt the largest possible reductions. Using a testbed in New York City, we monitored the electricity consumption in 36 apartments whose residents were sent 14 feedback messages over the course of 2 months. Using techniques akin to Natural Language Processing, the feedback messages were automatically generated to comprise 10 features in random combinations (e.g., with or without graph, with or without peer comparison). As a novelty vis-à-vis previous studies, each resident was sent different message types over the course of the experiment (instead of pre-partitioning the residents into sub-cohort who would henceforth receive the same message type throughout). In 504 observations, the feedback scheme prompted average reductions in electricity consumption of 11±3%, compared to 89 apartments which were monitored as a control group. Feedback elements that were particularly effective in prompting reductions were self-comparisons with one's own recent consumption (average 14%). Messages of high variety from one feedback to the next also prompted large reductions (average 16%). In contrast, feedback messages that included comparisons to neighbors did not prompt higher or lower reductions on average – but this average masked dispersion amongst residents: High baseline users tended to decrease their consumption in response to peer comparisons, and vice versa (boomerang effect). This behavior was consistent for all residents and could be explained by a simple mean reversion of each individual resident's use over time that was amplified by comparisons to neighbors. We discuss how the observed behavior represents norm-conformity rather than anti-conformity previously invoked in similar contexts. We further discuss how our findings can be used to optimize feedback messages in large scale field applications, for example with utilities.

Project: Multiple types of experiments over several years

Note: Following slides will focus only on one such exp.


Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Residential electricity conservation in response to auto-generated, multi-featured, personalized eco-feedback designed for large scale applications with utilities

Christoph J. Meinrenken[a,b,*], Sanjmeet Abrol[a], Gaurav B. Gite[a], Christopher Hidey[d], Kathleen McKeown[a,d], Ali Mehmani[a], Vijay Modi[a,c], Elsbeth C. Turcan[d], Wanlin Xie[d], Patricia J. Culligan[a,b,e,f]

Instead of sending an expert to someone's home ... use a low-cost, automated data-science approach:

Step 1: Measure apt. level loads at 10-sec. (real & reactive power) in ~400 apartments

Step 2: Break down to appliance level → Identify consumption hotspots (e.g., fridge)

Step 3: Determine other characteristics, e.g. phantom loads from electronic devices

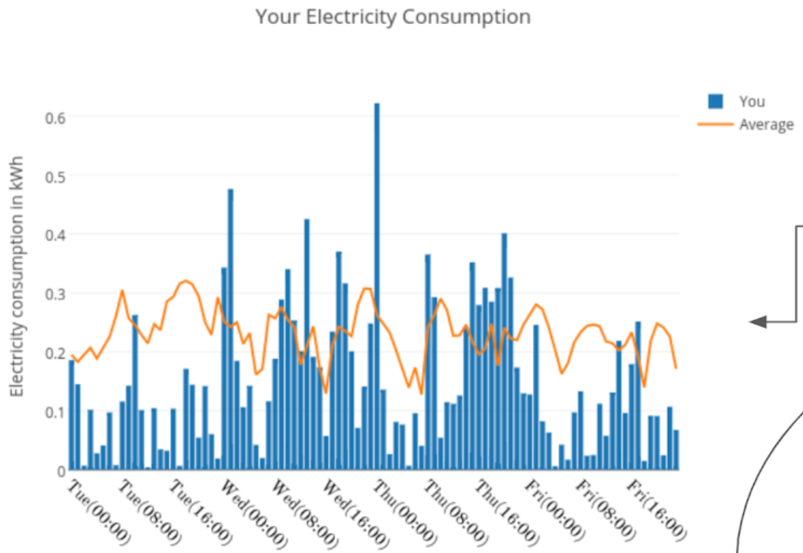

Image credit: n-architects

Step 6: Augment feedback messages with behavioral tips customized to their home, e.g.: "Did you know that cleaning the fridge grill can save substantial electricity." "Remember to turn off lights and unplug un-used electronics."

Step 5: Use NLP to mine online expert forums for electricity saving and load shifting tips

Step 4: Generate personalized feedback, e.g.: "Your fridge consumed 50% above average for your building." "Your electricity consumption never went below 120 Watt, causing $20 of your monthly bill."

Feedbacks comprised 10 randomly selected "features" …



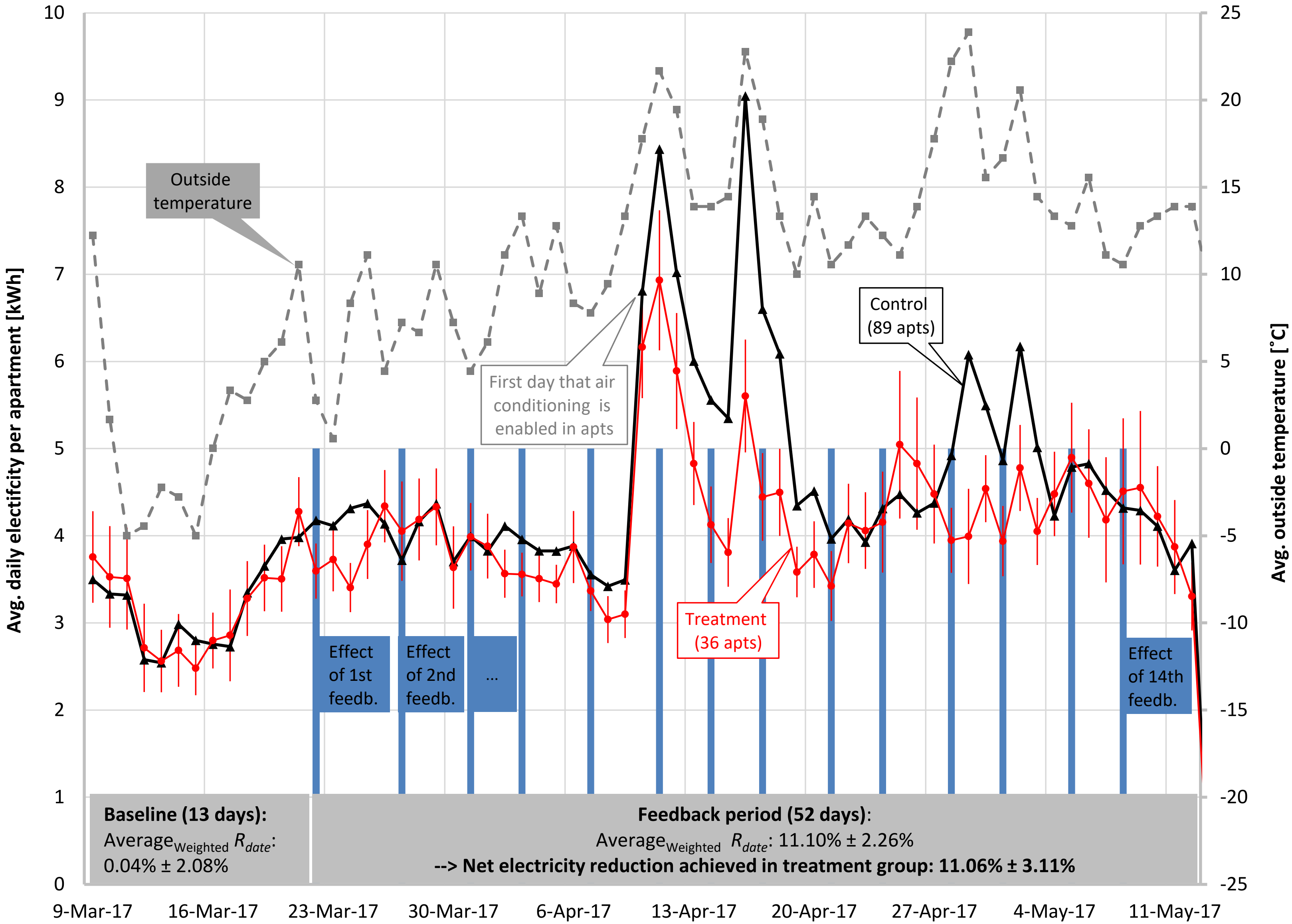
This feedback cycle (Monday-Friday), you used electricity that caused 4.33 kg of CO2 emissions. That's on a daily basis 22.04% less than your previous feedback cycle and 32.81% less than your peers' average, with the highest usage during the night-time. Great job! :)

For the month, you would be spending an estimated $5.02 less than your previous feedback cycle.

Sentiment - Positive
Equivalent cons. unit – CO_2 Emissions
Graph - True
Money Spent - False
Comparison With Peers (power) - True
Comparison With Self (power) - True
Comparison With Peers ($) - False
Comparison With Self ($) - True
Peak time information - True
News Articles - False

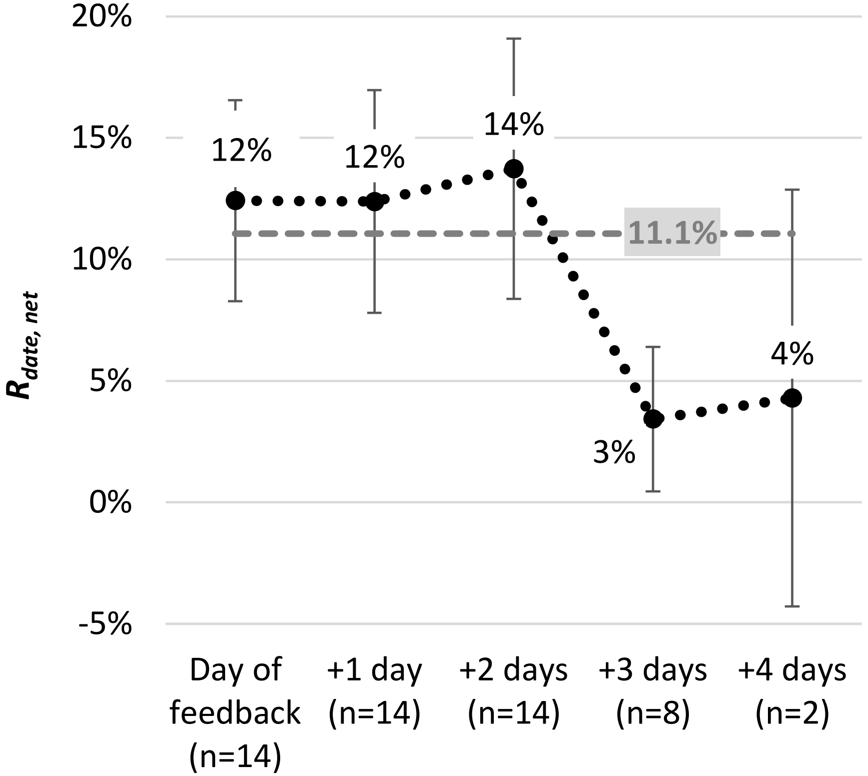
… so that the same resident received different message types from one feedback round to the next

We sent 14 feedbacks to 36 residents each ... and their electricity use dropped 11% vs. a control group



In-treatment persistence: Response relapsed substantially from third day on …

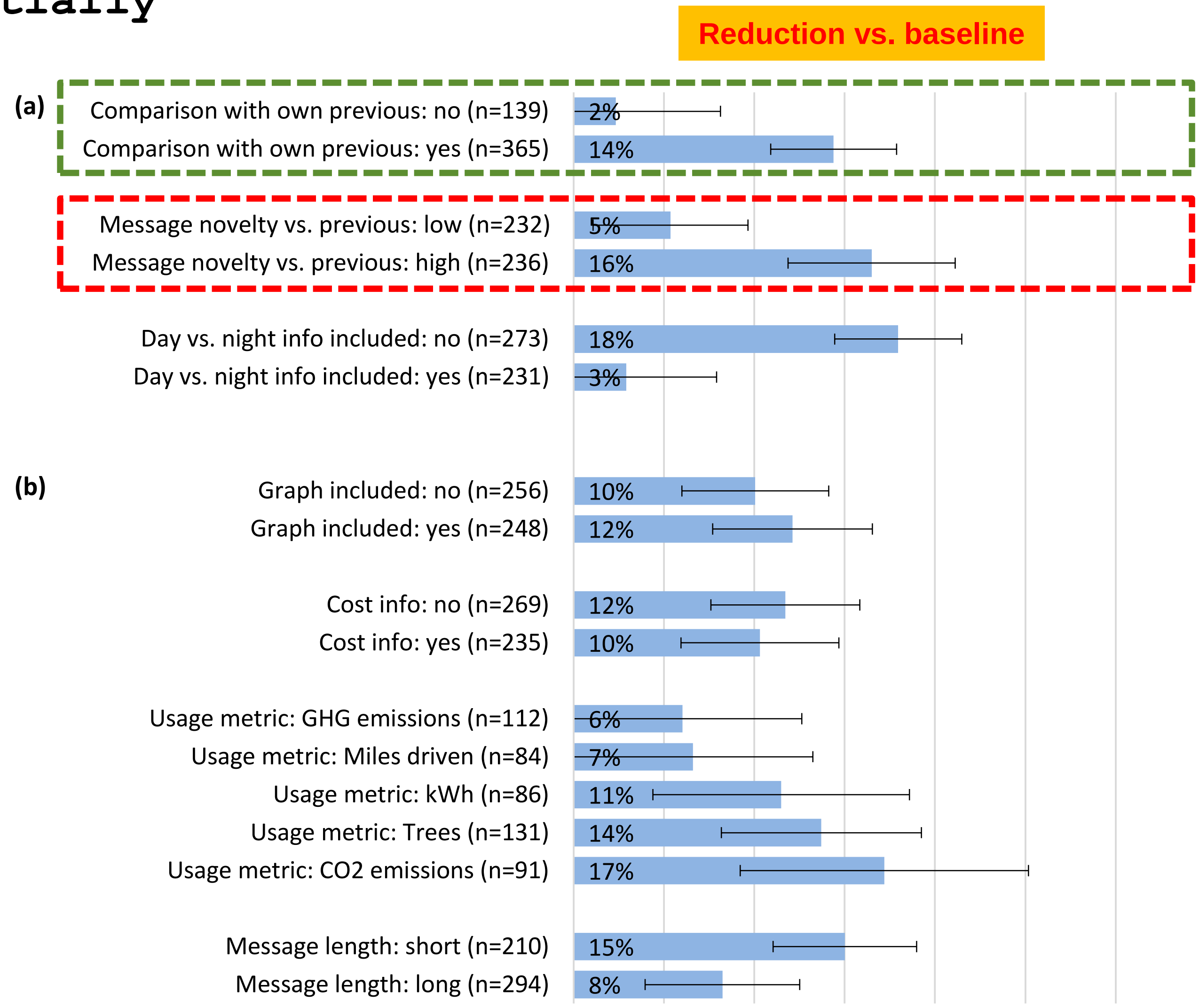
Figure 1. **Response-relapse.** Electricity usage reductions on specific days ($R_{date, net}$; 0-24h local time; see *Methods*) averaged across all 36 residents in the treatment group as a function of the time passed since the most recent feedback (feedbacks were sent at 10 a.m.). Error bars show ±1 SEM, accounting for the varying sample size. The reduction for the "+3 days" group is statistically significantly lower than for the 3 earlier days combined ($p<0.05$). Grey dashed line shows average $R_{date, net}$ across all 52 days of the feedback period.



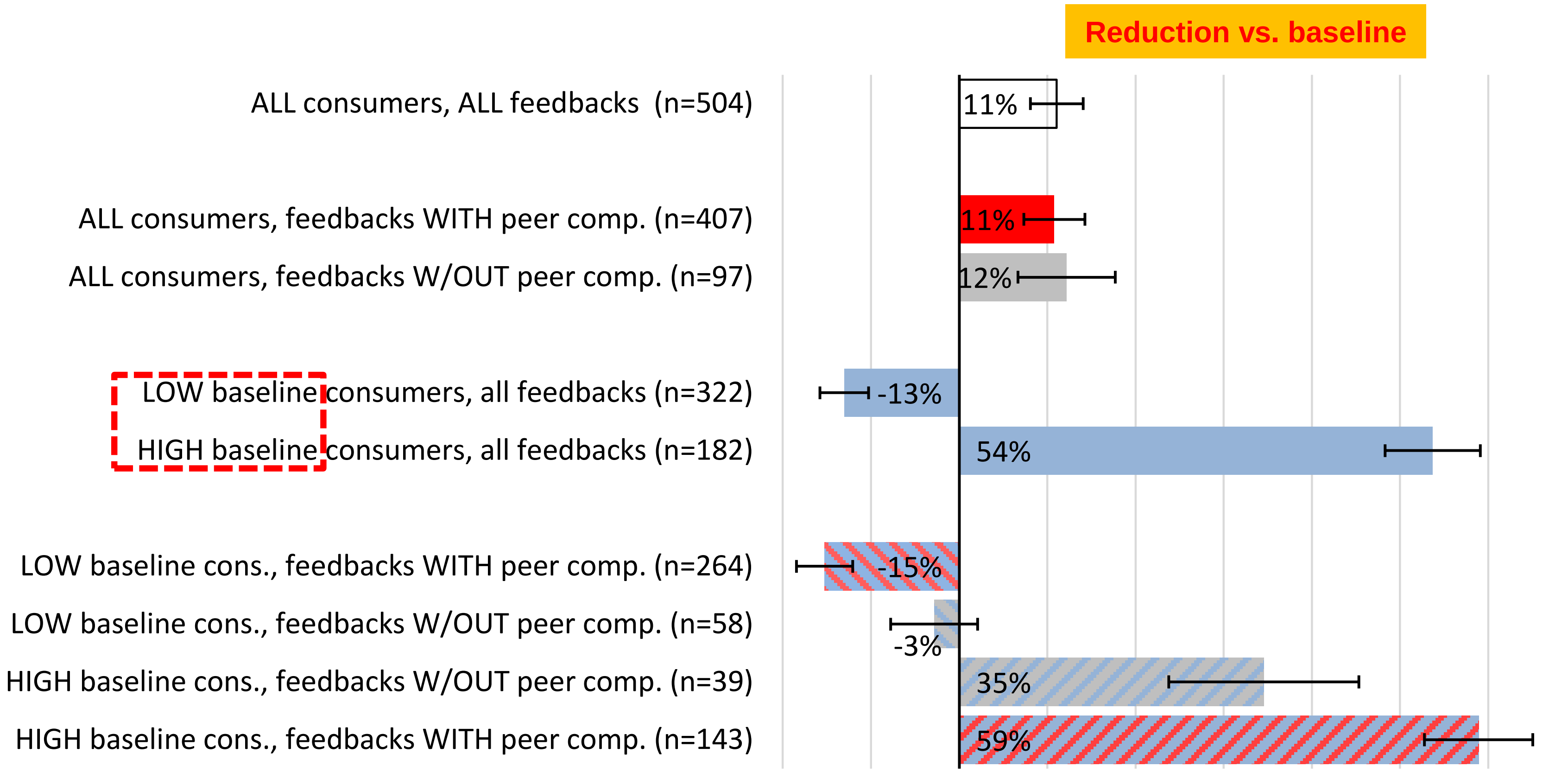
… yielding a dataset of 36*14=504 independent observations

Self-comparisons and message variation improved feedback efficacy substantially

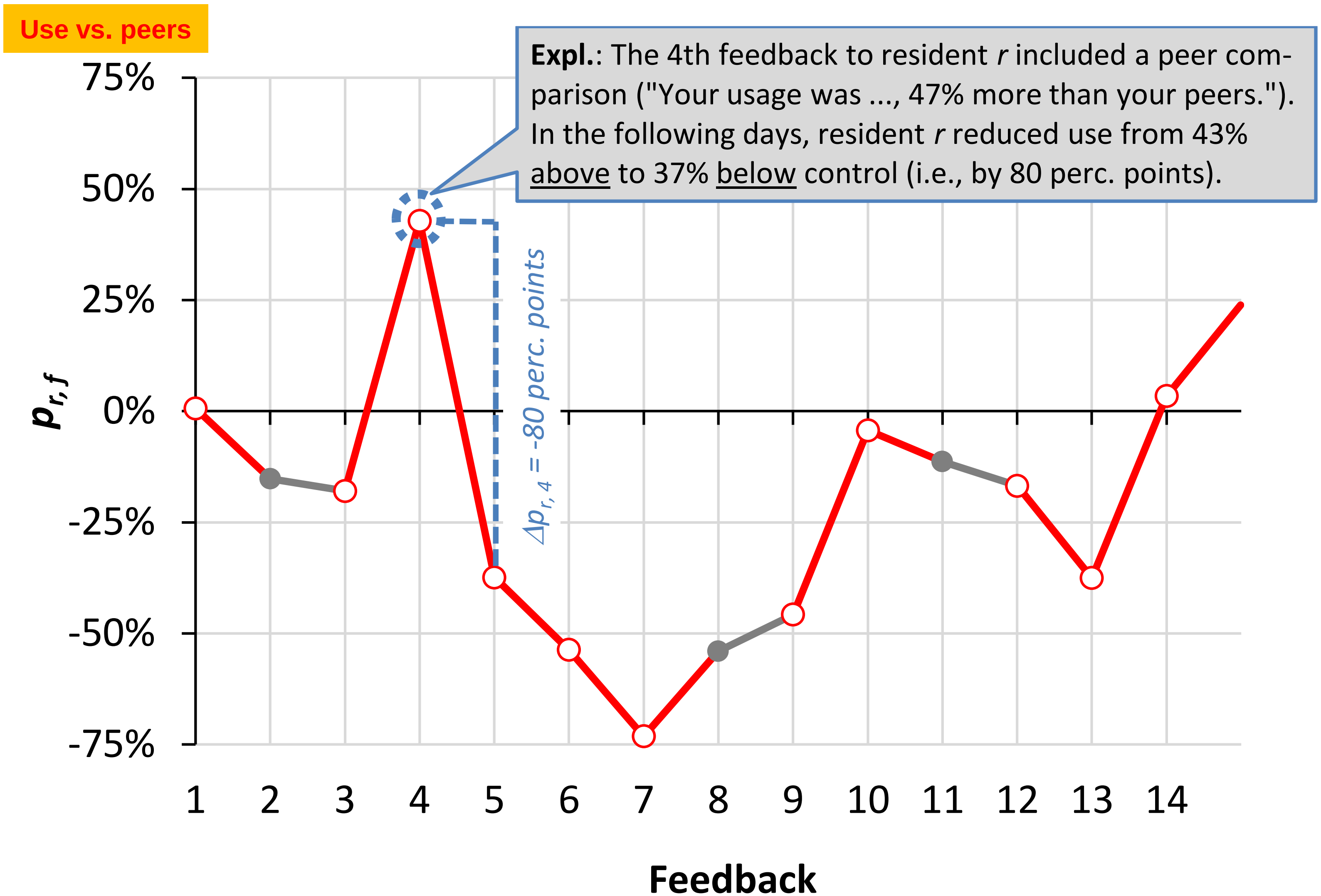
Figure 1. **Efficacy of feedback types.** Average electricity usage reductions as % of the control group ($R_{r,f}$; see *Methods*) as function of feedback type (average across all 504 $R_{r,f}$ is 11.1%). Error bars show ±1 SEM, accounting for the varying sample size. Results are shown in two tiers: (a) Types that had a (weakly) statistically significant effect on the observed reductions (p<0.10); (b) Types whose effect was directionally as expected based on previous literature, however not statistically significant given the sample size.



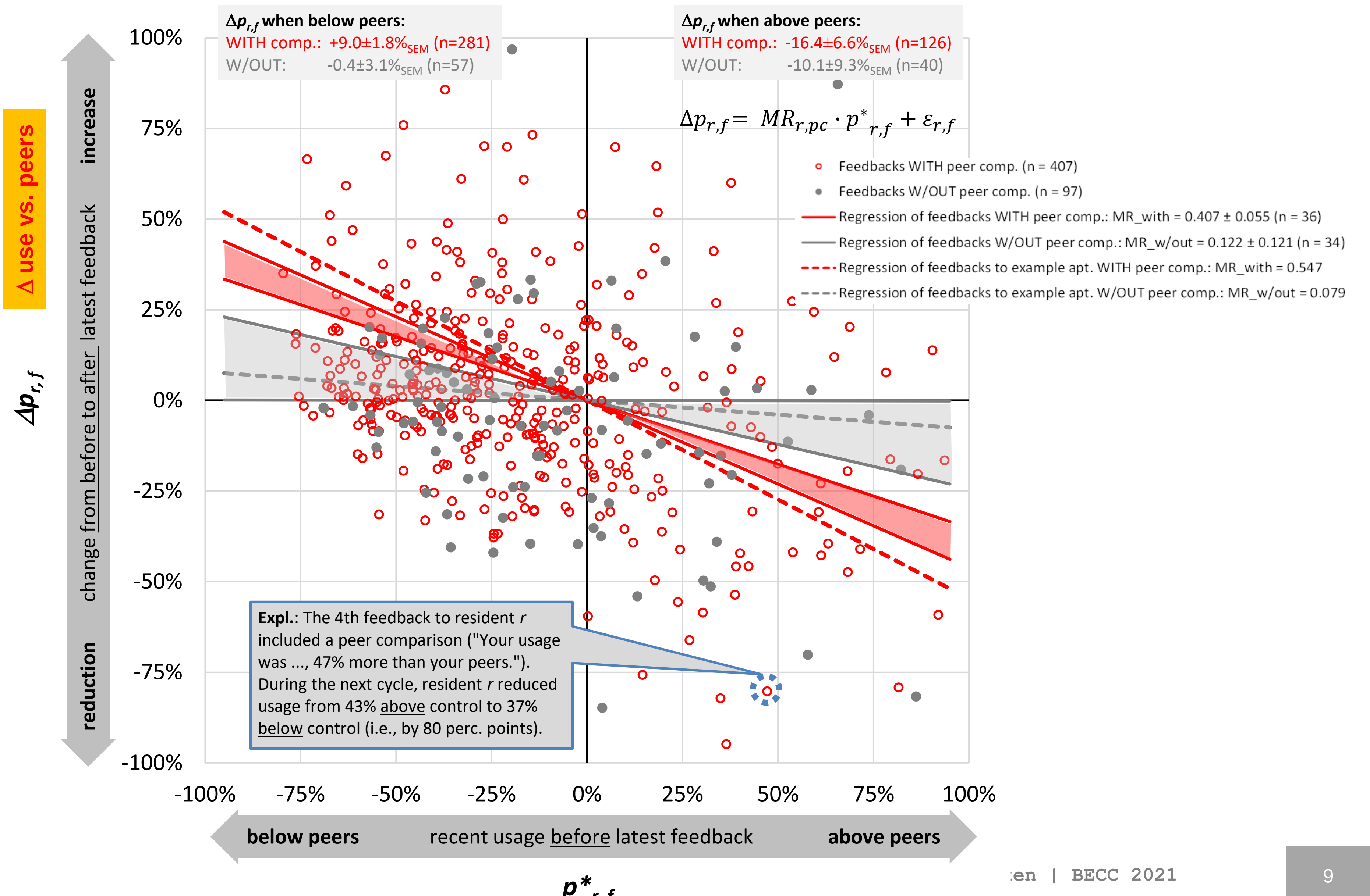
… **but the effect of peer comparisons was more complex,**
(consistent with previously observed boomerang effect)



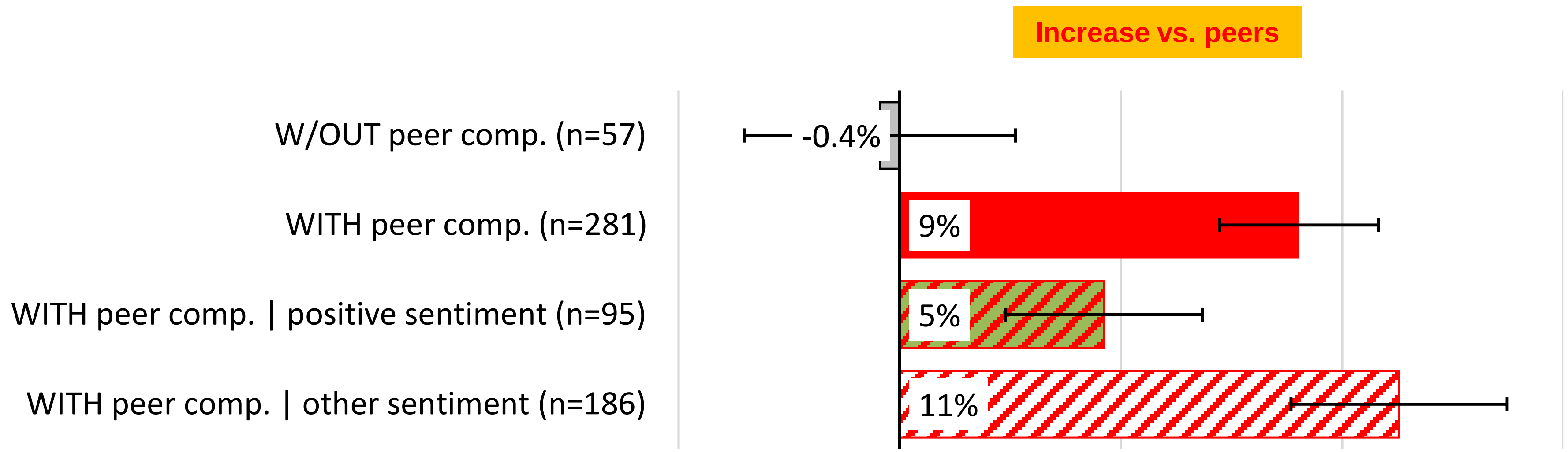
In a departure from previous work, we then tracked each resident's usage as a function of his/her most recent usage (instead of vs. their baseline)



Paired t-test for 36 residents showed that peer-comparisons amplified a naturally expected mean reversion effect ($p<0.05$)



Focusing on (recent) low consumers, positive sentiment mitigated the boomerang effect …



… but simply avoiding the peer comparison was more effective still

Summary and discussion

- 11.1±3.1% in line with Delmas et al. 2013 meta study
  - But what about 1.4-3.3% in O-Power experiments (Allcott 2011)?
  - Our 11% was measured only for the 42% residents who had opted in
    - (11.1±3.1%)·42% = 4.6±1.3%

Because of the specific experimental design of our study, novel results:

- Strategies for large scale field applications
  - Feedbacks with deliberate variation from one feedback to next
  - For residents with low <u>recent</u> usage … simply skip peer comparison
- Is it an anti-conform effect OR simply conformity?
  - Brehm and Brehm's (1981) have been summarized as: *"psychological reactance – that people act to protect their sense of freedom – is supported by experiments showing that attempts to restrict a person's freedom often produce an anti-conformity 'boomerang effect''"* (Myer's 2010)
  - Wesley Schultz describes it as a "magnet" effect to <u>both</u> sides
  - However, our observations are remarkably well explained by a simple random walk with mean reversion … implying that the observed boomerang effect was NOT born out of anti-conformity or even defiance. Instead: Every study participants simply wanted to conform to the average
  - Average acted like a "magnet", as postulated by e.g. Schultz et al, Psych. Sci. 18 (2007)