Planning to Save Energy: How Information Format Affects Accuracy

Thomas E. Gorman

Torsten Reimer

Juan Pablo Loaiza Ramirez

Hayden Barber

2024-12-07

Abstract

Effective communication of energy consumption information is crucial for promoting residential energy conservation. This study investigates how different numerical representations of energy reduction goals influence consumers’ ability to create accurate conservation plans. Across two experiments, we examined the impact of presenting energy information in kilowatt-hours (kWh), percentages, or U.S. dollars (USD) on planning accuracy. Participants completed a simulated household planning task in which they allocated energy usage across multiple appliances, with the goal presented in either kilowatt-hours (kWh), percentages, or monetary costs. Results across both experiments showed that presenting reduction goals in absolute units (kWh) led to significantly greater accuracy compared to percentage-based or monetary formats. Furthermore, we found that higher energy literacy was associated with more accurate planning. These findings demonstrate that absolute units (kWh) are more effective for communicating energy-saving goals, and highlight the potential value of educational interventions to improve consumer energy literacy.

# Introduction

### Literature Review

Energy costs often impose a significant burden on low-income households, leading to “energy insecurity,” where basic energy needs cannot be met (Bednar & Reames, 2020). Frequently measured as the energy burden—the percentage of income spent on energy bills—this burden can be disproportionately high for vulnerable families, necessitating difficult trade-offs with essentials such as food or medicine (Bednar & Reames, 2020; Memmott et al., 2021). Energy insecurity has been linked to health risks and unsafe coping strategies, disproportionately impacting racial and ethnic minorities through higher rates of disconnection (Memmott et al., 2021). While the broader context of climate change, partly driven by residential consumption (Farghali et al., 2023), underscores the need for sustainable solutions, the financial strain on vulnerable households remains a pressing concern. Promoting behavior change to reduce energy consumption is crucial. However, the success of such interventions may hinge on how effectively energy information is communicated, with format and presentation context significantly influencing understanding and action (Canfield et al., 2017; Fischer, 2008).

The way numerical information is presented can significantly affect how individuals process and use that information. Of particular relevance are reference class effects, which occur when numerical statements are presented without a clear or intuitive basis for comparison, thereby hindering the meaningful inference of quantities (Gigerenzer & Edwards, 2003; Reimer et al., 2015). A substantial body of evidence suggests that presenting data in terms of absolute counts or frequencies, as opposed to probabilities or percentages, can promote more accurate comprehension and facilitate. However, it’s important to note that even intuitive formats can pose challenges. For instance, Weber et al. (2018) found that individuals often struggle with reasoning tasks presented in natural frequencies because they inadvertently revert to more complex probabilistic thinking.

The choice of units and the format in which information is presented have also been shown to exert a significant influence on decision-making specifically within the context of energy consumption and planning. For instance, tables are generally more effective than graphs for conveying specific electricity usage data because they facilitate straightforward point reading Canfield et al. (2017). Furthermore, the framing of energy costs, such as displaying monthly rather than daily or yearly expenses, can significantly affect consumers’ choices (Gill et al., 2022). Consequently, the selection of an appropriate information format is crucial for effectively supporting energy-related decisions. Therefore, similar to natural frequencies, kWh provide a direct measure of energy use, a characteristic that could simplify calculations and facilitate comparisons, potentially helping consumers better understand and compare the energy consumption of different appliances or activities. Conversely, prior research has suggested that consumers have a preference for receiving energy feedback in terms of monetary values over scientific units (Karjalainen, 2011; Nemati & Penn, 2020), as well as better long-term appliance selection with information presented in monetary terms (Blasch et al., 2019). However, it remains uncertain how these reported benefits might generalize to the more complex domain of household energy planning, where usage patterns can be multifaceted and subject to a variety of contextual influences.

Although many individuals express a desire to conserve energy, research consistently shows that abstract goals (e.g., “reduce overall usage by 15%”) often fail to translate into effective behavior change unless accompanied by specific, actionable steps (Abrahamse et al., 2005; Nemati & Penn, 2020). For instance, Abrahamse et al. (2005) demonstrated that merely providing general information about energy savings rarely alters consumption patterns unless consumers also receive concrete instructions or tailored feedback. Similarly, Tonke (2024) reported that sending households brief but precise text messages outlining how to reduce water use (e.g., limiting irrigation times, adjusting washing machine settings) yielded meaningful decreases in consumption, underscoring the importance of procedural knowledge—namely, knowing how to operationalize a goal rather than simply why it is desirable. In the context of energy conservation, this implies that interventions should not only highlight potential reductions (such as a 15% target) but also guide residents in allocating those reductions across specific appliances or behaviors (Attari et al., 2010). Additionally, meta-analytic findings suggest that people respond more robustly to household-level feedback that situates their usage within a personalized framework, thereby reducing the cognitive burden of figuring out next steps (Nemati & Penn, 2020).

Despite existing studies on energy-use communication and format effects, limited research has explored how different numerical representations influence consumers’ ability to create accurate energy conservation plans. Specifically, there is a gap in understanding how presenting energy information in absolute units versus percentages or monetary terms affects the precision of planning appliance-specific reductions. The current study addresses these critical issues by systematically investigating the impact of varying information formats (kWh, percentage, and USD) on the accuracy of energy-planning decisions. By manipulating the presentation format of energy information, this research aims to elucidate how different representational formats influence planning accuracy.Based on the literature reviewed, we hypothesize that: 1) Presenting energy reduction goals in absolute units (kWh) will lead to greater planning accuracy compared to percentage-based or monetary formats, as absolute units provide a more direct and less ambiguous representation of energy quantities. 2) Higher energy literacy will be associated with more accurate planning, as individuals with greater energy knowledge may be better equipped to process and utilize the provided information, regardless of format. We also examine the potential of several exploratory variables, such as goal difficulty and the rounding of numerical values, to further elucidate the factors that influence planning accuracy.

# Experiment 1

See [Figure 1](#fig-task) for an example of a planning trial as it was seen by participants.

## Methods

### Participants

We implemented our task and surveys on Qualtrics, and recruited participants through Amazon Mechanical Turk. In Experiment 1, 252 participants were initially recruited, but data from 17 participants were corrupted due to experimenter error, leaving a final sample of 235 participants. Most participants (76%) reported using a calculator to complete the task.

### Materials and Design

The study employed a mixed design with reference class (kWh, percentage, USD) as a between-subjects factor and state/family scenario as a within-subjects factor. Each participant completed energy reduction planning tasks for two different states, with state order counterbalanced across participants. The family scenarios featured four households in different climate regions: Texas (Smith family) and California (Adams family) representing warm climates, and Colorado (Wells family) and Massachusetts (Davis family) representing cold climates. We obtain average utility use from each state by CITE SOURCE FOR STATE AVGS?

### Procedure

Participants were provided with energy usage data for two hypothetical families and tasked with creating action plans to meet specified energy reduction goals. These goals were implementing by allocating usage across five appliance categories: heating, cooling, water heating, refrigerator, and other appliances (e.g., TV, lighting).

For each family scenario, participants were shown a table containing the families utility usage from the prior year, alongside the state averages for each appliance category (both prior year usage and state averages are always shown in kWh). For each scenario, participants were asked to create two possible action plans to achieve the target reduction in total household energy usage (see [Figure 1](#fig-task)). Depending on their reference class condition, the target reduction amount presented either in kilowatt-hours (kWh), as percentages of total household usage, or in U.S. dollars. In all conditions, the target reduction was equivalent to a 15% reduction in total household kWh.

kWh Condition: The target reduction is presented as an absolute number of kilowatt-hours (e.g., “Reduce total energy use by 5965 kWh”).

Percentage Condition: The target reduction is presented as a percentage of the family’s total previous year’s usage (e.g., “Reduce total energy use by 15%”).

USD Condition: The target reduction is presented as a monetary value, equivalent to the cost of a 15% reduction in energy use (e.g., “Reduce total energy costs by $656”).

|  |
| --- |
| Figure 1: Example energy planning task trial. Participants saw a table with a family’s previous year electricity usage (here for the Wells family in Colorado) and were asked to allocate energy usage to meet a 15% reduction goal. The format of the reduction goal was manipulated to be either a percentage (15% given as goal reduction), kilowatt hours (5965 kWh given), or USD ($656) |

Additional data collected included:

* **Energy Literacy Quiz**: An 8-item questionnaire assessing participants’ knowledge of energy consumption and conversion (DeWaters & Powers, 2011).
* **Calculator Usage Tracking**: Questions determined whether participants used a calculator, paper/pen, or other methods to complete the tasks.

## Results

### Data Analysis

All preprocessing and analyses were carried out in R (Team, 2020) and the tidyverse package (Wickham et al., 2019). Mixed Bayesian regressions were fit using the brms package (Bürkner, 2017), with participants and family scenario (states) set as random effects.

pacman::p\_load(dplyr,purrr,tidyr,stringr,here,tibble,brms,rstan,bayestestR,emmeans,tidybayes,  
 ggplot2,gt,knitr,kableExtra,ggh4x,patchwork, ggridges,ggstance,lme4,flextable,pander,marginaleffects)  
  
options(digits=2, scipen=999, dplyr.summarise.inform=FALSE)  
  
walk(c("fun\_plot"), ~ source(here::here(paste0("scripts/", .x, ".R"))))  
  
theme\_set(theme\_nice())  
  
s1 <- readRDS(here::here("data/s1\_processed.rds")) |>   
 filter(!(id %in% readRDS(here::here("data/s1\_discrep\_ids.rds")))) |>   
 filter(!(id %in% readRDS(here::here("data/s1\_grp\_outlier\_ids.rds")))) |>  
 mutate(refClass = factor(refClass, levels=c("kWh","Percentage","USD")))  
  
s2\_long <- readRDS(here::here("data/s2\_processed.rds")) |>   
 filter(!(id %in% readRDS(here::here("data/s2\_discrep\_ids.rds")))) |>   
 filter(!(id %in% readRDS(here::here("data/s2\_grp\_outlier\_ids.rds"))) ) |>  
 mutate(refClass = factor(refClass, levels=c("kWh","Percentage","USD")))

s1\_agg <- s1 |>   
 filter(appliance !="Total kWh") |>   
 group\_by(id,refClass,state,block,plan,edu,pct\_goal,calc) |>   
 summarise(total\_kWh = sum(value),orig\_kWh=sum(family),   
 pct\_change = abs(round((orig\_kWh-total\_kWh)/orig\_kWh,3)),   
 n\_change = sum(value!=family),  
 state\_p\_dif=mean(state\_p\_dif),  
 state\_f\_dif=mean(state\_f\_dif),  
 n\_less\_avg = sum(less\_avg),  
 duration=first(duration), .groups = 'drop') |>   
 mutate(matched\_goal = (pct\_change == pct\_goal),   
 error = pct\_change - pct\_goal,  
 abs\_error = abs(error),  
 log\_abs\_error=log(abs(error)+.007),   
 close\_match = abs\_error <= 0.02) |>  
 mutate(  
 accuracy\_level = factor(  
 case\_when(  
 abs\_error == 0.00 ~ "Exact match",  
 abs\_error <= 0.05 ~ "0.01-5% error",  
 TRUE ~ "Over 5% error"   
 ),   
 levels = c("Exact match", "0.01-5% error", "Over 5% error"),  
 ordered = TRUE  
 )  
 ) |> relocate(accuracy\_level, .after= "pct\_change")  
  
  
  
  
s1\_agg4 <- s1\_agg |> group\_by(id,refClass,calc) |>   
 mutate(n\_accuracy = n\_distinct(accuracy\_level)) |>   
 summarise(  
 mg=sum(matched\_goal),  
 mgc=sum(close\_match),  
 n=n(),   
 pct=mg/n,  
 pct\_close=mgc/n,  
 mean\_pct\_change=mean(pct\_change),  
 mean\_abs\_error=mean(abs\_error),  
 mean\_log\_abs\_error=mean(log\_abs\_error)) |>   
 mutate(accuracy\_level = factor(  
 case\_when(  
 mean\_abs\_error < 0.02 ~ "Exact match",  
 mean\_abs\_error <= 0.05 ~ "0.01-5% error",  
 TRUE ~ "Over 5% error"   
 ),   
 levels = c("Exact match", ".01-5% error", "Over 5% error"),  
 ordered = TRUE  
 ))

##| label: tbl-s1-agg  
##| tbl-cap: "Study 1: Summary of planning accuracy by reference class. The table shows performance as both the % of trials where participants matched the goal, and the mean absolute error from the target reduction goal."  
  
# overall pct of subjects who matched their goal  
s1\_agg4 |> group\_by('Reference Class' = refClass) |>  
 summarise(  
 'Avg. % Change' = mean(mean\_pct\_change),  
 '% meeting goal (exact)' = mean(pct),  
 '% meeting goal (close match)' = mean(pct\_close),  
 'Abs. Deviation' = median(mean\_abs\_error),  
 'Log Abs. Deviation' = (median(mean\_log\_abs\_error)),  
 # sd = sd(pct),  
 # n = n(),  
 #se=sd(pct)/sqrt(n)  
) |> mutate(across(where(is.numeric), \(x) round(x, 3))) %>%   
 kable(escape=FALSE,booktabs=TRUE,align=c("l"))   
  
#pander::pandoc.table(caption="Study 1: Proportion of participants who matched their goal overall")

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Study 1: Summary of planning accuracy by reference class. The table shows performance as both the % of trials where participants matched the goal, and the mean absolute error from the target reduction goal   | Reference Class | Avg. % Change | % meeting goal (exact) | % meeting goal (close match) | Abs. Deviation | Log Abs. Deviation | | --- | --- | --- | --- | --- | --- | | kWh | 0.22 | 0.38 | 0.54 | 0.03 | -3.7 | | Percentage | 0.21 | 0.22 | 0.40 | 0.06 | -3.1 | | USD | 0.23 | 0.10 | 0.22 | 0.10 | -2.4 | |

[Table 1](#tbl-s1-agg) that participants in the kWh condition met the target goal 38% of the time, compared to 22% for the Percentage condition and 10% for the USD condition. Moreover, the kWh reference class exhibited smaller deviations from the target reduction, suggesting that participants performed more accurately when the goal was framed in kWh rather than when percentages or USD.

As shown in [Table 1](#tbl-s1-agg), participants in the kWh condition exactly met the target reduction goal 38% of the time, significantly outperforming those in the Percentage (22%) and USD (10%) conditions. Furthermore, the kWh reference class exhibited notably smaller mean absolute deviations (0.03) compared to Percentage (0.06) and USD (0.10), suggesting that presenting the reduction goal in absolute units facilitated more precise allocations.

s1\_ld <- ggplot(s1\_agg, aes(y = refClass, x = log\_abs\_error, fill = refClass)) +  
 geom\_density\_ridges(aes(col = refClass), alpha = 0.2, scale = 0.5,  
 jittered\_points = TRUE, point\_alpha = 0.7,point\_size=.4,  
 position = position\_raincloud(width = 0.05, height = 0.1,  
 ygap = 0.05)) +  
 geom\_boxploth(width = 0.1, alpha = 0.2, outlier.shape = NA, show.legend = FALSE) +  
 #scale\_y\_discrete(expand = expansion(mult = c(0.2, 0.4))) +  
 # guides(fill = "none", color = guide\_legend(reverse = TRUE)) +  
 guides(fill = "none") +  
 labs(x = "Log Absolute Deviation", y = "Reference Class", color = "Reference Class") +  
 theme(legend.position = "top")  
  
  
s1\_ldc <- ggplot(s1\_agg, aes(y = refClass, x = log\_abs\_error, fill = calc)) +  
 geom\_density\_ridges(aes(col = calc), alpha = 0.2, scale = 0.5,  
 jittered\_points = TRUE, point\_alpha = 0.7, point\_size = .4,  
 position = position\_raincloud(width = 0.05, height = 0.1,  
 ygap = 0.05)) +  
 geom\_boxploth(width = 0.1, alpha = 0.2, outlier.shape = NA, show.legend = FALSE) +  
 scale\_color\_brewer(palette = "Set1") +  
 scale\_fill\_brewer(palette = "Set1") +  
 guides(fill = "none") +  
 labs(x = "Log Absolute Deviation", y = "Reference Class", color = "") +  
 theme(legend.position = "top")  
  
s1\_ld | s1\_ldc

|  |
| --- |
| Figure 2: Experiment 1: Distribution of the log of the absolute error between the participant’s action plan and the reduction goal across different reference class conditions (kWh, Percentage, USD). The right side plots are further separated by calculator usage. A lower log absolute error suggests higher planning accuracy. |

# compute percentage of subjects per accuracy level per group  
observed\_props\_s1 <- s1\_agg |>  
 group\_by(refClass, accuracy\_level) |>  
 summarise(n = n()) |>  
 group\_by(refClass) |>  
 mutate(prop = n/sum(n)) |>  
 mutate(n\_prop=paste0(n," (",round(prop\*100,1),"%)" ), pct\_grp=paste0(round(prop\*100,1), "%")) |> ungroup()  
  
observed\_props\_s1 |>   
 mutate(n\_total=sum(n)/4) |>   
 group\_by(accuracy\_level) |>  
 mutate(ns=sum(n)/4) |>   
 mutate(Total = paste0(round(ns/n\_total\*100,1), "%")) |>  
 select('Reference Class'=refClass, 'Accuracy Level'=accuracy\_level, '% in Group'=pct\_grp, "Combined Groups %" =Total) |>  
 pivot\_wider(  
 names\_from = 'Reference Class',  
 values\_from = c('% in Group')  
 ) |> relocate("Combined Groups %" , .after=last\_col()) |>   
 kable(escape=FALSE,booktabs=TRUE,align=c("l"))

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Study 1: The table shows the percentage of participants who fell into each accuracy level for each reference class condition (percentages of kWh, $, and USD columns reflect within condition percentages). The combined group column reflects the percentage of participants in each accuracy level when aggregating across across all reference class conditions.   | Accuracy Level | kWh | Percentage | USD | Combined Groups % | | --- | --- | --- | --- | --- | | Exact match | 38.5% | 22.4% | 10.2% | 23.1% | | 0.01-5% error | 22.7% | 29.5% | 25% | 25.5% | | Over 5% error | 38.8% | 48.1% | 64.8% | 51.3% | |

We next categorized responses into three accuracy levels (exact match [0% error], minor deviations [0.01–5%], and large deviations [>5%]) for our primary statistical modeling. Using Bayesian ordinal regression, we modeled the ordered accuracy outcome as a function of the reference class condition, while controlling for random variation across participants and family scenarios:

This approach allowed us to estimate thresholds (intercepts) and regression coefficients that capture how different reference classes affect the likelihood of achieving higher accuracy categories. For each comparison, we provide posterior odds ratios (OR) and their 95% CIs. This approach allows the estimation of threshold parameters and regression coefficients that characterize how changes in predictor variables (such as the reference class: kWh, percentage, or USD) relate to probabilities of being in each accuracy category. Specifically, we used a cumulative logit link function to model the ordered accuracy outcome, and we specified weakly informative priors for the regression coefficients (normal distributions with mean 0 and standard deviation of 1) and for the cutpoints (normal distributions with a mean of zero and a standard deviation of 4.0). The approach allows us to estimate threshold parameters and regression coefficients that characterize how changes in predictor variables (such as the reference class: kWh, percentage, or USD) relate to probabilities of being in each accuracy category.

##| label: tbl-s1-ord  
##| tbl-cap: "Study 1: Ordinal Regression Model Results."  
  
  
ordinal\_model\_s1 <- brm(  
 accuracy\_level ~ refClass +calc + (1|id) + (1|state),  
 data = s1\_agg,  
 family = cumulative("logit"),  
 cores = 4,  
 iter = 4000,  
 control = list(adapt\_delta = 0.98),   
 prior = c(prior(normal(0, 4), class = "Intercept"),   
 prior(normal(0, 4), class = "b")),   
 file = paste0(here::here("data/model\_cache",'s1\_acc3\_add.rds'))   
)  
  
t1 <- as.data.frame(describe\_posterior(ordinal\_model\_s1, centrality = "Mean"))[, c(1,2,4,5,6)] |>   
 setNames(c("Parameter", "Estimate", "CI\_Lower", "CI\_Upper", "pd")) |>   
 mutate(Parameter = stringr::str\_remove(Parameter, "b\_")) |> kable(escape=FALSE,booktabs=TRUE,align=c("l"), row.names = FALSE)  
  
# Get predicted probabilities  
# pred\_summary <- ordinal\_model\_s1 |>  
# epred\_draws(newdata = data.frame(refClass = c("kWh", "Percentage", "USD")),  
# ndraws = 1000, re\_formula = NA) |>  
# group\_by(refClass, Category=.category) |>  
# summarise(  
# mean\_prob = mean(.epred),  
# lower\_ci = quantile(.epred, 0.025),  
# upper\_ci = quantile(.epred, 0.975)  
# )  
#pred\_summary |> pander::pandoc.table(caption="Study 1: Predicted probabilities of accuracy")  
  
#odds ratios of fixed effects  
# as.data.frame(fixef(ordinal\_model\_s1)[,-2])|> as.data.frame() %>%  
# rownames\_to\_column(var = "Parameter") %>%  
# mutate(across(where(is.numeric), exp)) |>  
# filter(!stringr::str\_detect(Parameter, "Intercept")) |>   
# filter(!stringr::str\_detect(Parameter, "calc")) |>   
# # rename columns to |comparison | odds\_ratio| ci\_lower| ci\_upper|  
# rename(Comparison = Parameter, odds\_ratio = Estimate, ci\_lower = Q2.5, ci\_upper = Q97.5) |>  
# # rename levels to Percentage vs kWh; USD vs kWh;   
# mutate(Comparison = case\_when(  
# str\_detect(Comparison, "refClassPercentage") ~ "Percentage vs kWh",  
# str\_detect(Comparison, "refClassUSD") ~ "USD vs kWh",  
# TRUE ~ Comparison  
# )) |> kable(escape=FALSE,booktabs=TRUE,align=c("l"))

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: **Experiment 1**: Ordinal Regression results. Ordinal regression results. Positive coefficients for the reference class predictors indicate that those conditions are associated with higher error categories relative to the kWh baseline.   | Parameter | Estimate | CI\_Lower | CI\_Upper | pd | | --- | --- | --- | --- | --- | | Intercept[1] | -4.21 | -5.90 | -2.58 | 1.00 | | Intercept[2] | -0.89 | -2.49 | 0.71 | 0.87 | | refClassPercentage | 1.44 | 0.07 | 2.88 | 0.98 | | refClassUSD | 3.13 | 1.81 | 4.50 | 1.00 | | calcUsedCalculator | -3.30 | -4.80 | -1.92 | 1.00 | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4: **Experiment 1**: Odds ratios for group comparisons. Odds ratios greater than 1 indicate increased odds of falling into a worse accuracy category compared to the kWh condition.   | Comparison | odds\_ratio | ci\_lower | ci\_upper | | --- | --- | --- | --- | | Percentage vs kWh | 4.2 | 1.1 | 18 | | USD vs kWh | 22.9 | 6.1 | 90 | |

As shown in [Table 3](#tbl-s1-reg), the reference class coefficients are positive for both the Percentage (Estimate = 1.3, 95% CI: 0.01 to 2.66, pd = 0.98) and USD (Estimate = 2.8, 95% CI: 1.52 to 4.04, pd = 1.00) conditions, relative to the kWh baseline. This indicates that, compared to the kWh condition, participants in both the Percentage and USD conditions were more likely to produce plans that fell into higher error categories. Moreover, the odds ratios (see Table 3) suggest that the USD condition led to a notably higher likelihood of large errors compared to the kWh baseline (OR = 15.7), while the Percentage condition also demonstrated increased odds (OR = 3.7) but was somewhat less detrimental to accuracy than USD. These results align with our descriptive findings and further clarify that framing the target reductions in absolute kWh units may facilitate significantly more accurate planning. Posterior predictive checks showed that the ordinal model provided a reasonable fit to the observed data (see [Figure 3](#fig-s1-ppd)).

pp\_check(ordinal\_model\_s1, type = "bars\_grouped", group="refClass", fatten = 2) +  
 scale\_x\_continuous("Response Category", breaks = 1:3,   
 labels = c("Exact", "0.01-5%", ">5%")) +  
 scale\_y\_continuous(expand = expansion(mult = c(0, 0.05))) +  
 ggtitle("Posterior Predictive Check by Reference Class") +  
 theme\_minimal() +  
 theme(  
 legend.background = element\_blank(),  
 legend.position = "bottom",  
 panel.grid.minor = element\_blank(),  
 axis.text.x = element\_text(angle = 45, hjust = 1)  
 )

|  |
| --- |
| Figure 3: Experiment 1: Posterior predictive check for frequency of trials at each accuracy level, faceted by reference class (kWh, Percentage, USD). Bars show observed, dots show model predicted proportions. Better accuracy is indicated by higher proportions in the ‘Exact Match’ and ‘0.01-5% error’ categories. |

s1\_els\_log\_error <- brm(  
 log\_abs\_error ~ els + (1|id) + (1|state),  
 data = s1\_agg,  
 family = gaussian(),  
 cores = 4,  
 iter = 2000,  
 control = list(adapt\_delta = 0.97),   
 prior = c(prior(normal(0, 3), class = "Intercept"),   
 prior(normal(0, 3), class = "b")),   
 file = paste0(here::here("data/model\_cache",'s1\_els\_log\_error.rds'))   
)  
  
# summary(s1\_els\_log\_error)  
# conditional\_effects(s1\_els\_log\_error)  
  
  
# Create the conditional effects plot  
conditional\_effects\_plot <- conditional\_effects(s1\_els\_log\_error)  
  
# Extract the data for plotting  
plot\_data <- conditional\_effects\_plot[[1]]  
  
# Create the plot  
ggplot(plot\_data, aes(x = els, y = estimate\_\_)) +  
 geom\_line(color = "blue", size = 1) +  
 geom\_ribbon(aes(ymin = lower\_\_, ymax = upper\_\_), alpha = 0.2) +  
 labs(  
 x = "Energy Literacy Score",  
 y = "Log Absolute Error",  
 title = "Conditional Effect of Energy Literacy on Log Absolute Error"  
 ) +  
 theme\_minimal()

|  |
| --- |
| Figure 4: Experiment 1. Conditional effect of energy literacy on log absolute error. The plot shows the relationship between energy literacy score and log absolute error, controlling for random effects of participant and state. Higher energy literacy scores are associated with smaller deviations from the target reduction goal, indicating more accurate planning. |

To further investigate individual factors that may influence planning accuracy, we examined the relationship between participants’ energy literacy scores and their performance on the task. Energy literacy was assessed using an 8-item questionnaire adapted from (DeWaters & Powers, 2011), which covers topics such as energy units, appliance energy consumption, and sources of electricity. A Bayesian linear regression model was fit with log-transformed absolute error as the outcome variable and energy literacy score as the predictor, controlling for random effects of participant and state: log\_abs\_error ~ els + (1|id) + (1|state). Results indicated a significant negative relationship between energy literacy and log absolute error (Estimate = -2.35, 95% CI: -2.88 to -1.81), suggesting that participants with higher energy literacy scores tended to have smaller deviations from the target reduction goal, and thus more accurate plans overall ([Figure 4](#fig-s1-els)).

## Experiment 1: Discussion

Experiment 1 examined how different numerical representations of energy reduction goals affected participants’ planning accuracy. In line with our hypothesis that absolute units would yield better accuracy, the kWh condition supported significantly more precise energy reduction plans than did either the Percentage or USD conditions. Although the Percentage format was detrimental to accuracy relative to kWh, it was the USD condition that consistently produced the poorest outcomes, suggesting that monetary terms, while intuitive in everyday contexts, may not serve as effective reference classes for planning appliance-specific reductions in energy use.

Experiment 2 will extend these findings by examining whether additional variables, such as the difficulty of the reduction goal or the rounding of numerical values, further interact with reference class conditions, thereby providing a more comprehensive understanding of how to optimize energy information presentation for improved planning accuracy.

# Experiment 2

## Methods

The experimental procedures in Experiment 2 are quite similar to those in Experiment 1. Experiment 2 employed a 2 (task goal: 10% vs. 15% reduction) x 2 (last year’s usage: exact vs. rounded) within-subjects design, with a between-subjects manipulation of the reference class (USD vs. Percentage vs. kWh). We recruited 206 participants from Amazon Mechanical Turk, but data from from 10 participants were corrupted due to experimenter error, leaving a final sample of 196 participants.

## Results

s2\_agg <- s2\_long |>   
 filter(appliance != "TOTAL") |>   
 group\_by(id,refClass,calc, state,pct,pct\_goal,plan,rounded) |>   
 summarise(  
 total\_kWh = sum(value),  
 orig\_kWh = sum(family),  
 pct\_change = round((orig\_kWh - total\_kWh) / orig\_kWh, 3),  
 state\_dif = mean(state\_dif),  
 .groups = "drop"  
 ) |>  
 mutate(  
 matched\_goal = (pct\_change == pct),  
 close\_match = abs(pct\_change - pct) <= 0.02,  
 error = pct\_change - pct,  
 abs\_error = abs(error),  
 log\_abs\_error=log(abs(error)+.007)) |>   
 mutate(  
 accuracy\_level = factor(  
 case\_when(  
 abs\_error == 0.00 ~ "Exact match",  
 abs\_error <= 0.05 ~ "0.01-5% error",  
 TRUE ~ "Over 5% error" # Capture all remaining cases  
 ),   
 levels = c("Exact match","0.01-5% error", "Over 5% error"),  
 ordered = TRUE  
 )  
 )  
  
s2\_agg4 <- s2\_agg |> group\_by(id,refClass,calc) |>   
 mutate(n\_accuracy = n\_distinct(accuracy\_level)) |>   
 summarise(  
 mg=sum(matched\_goal),  
 mgc=sum(close\_match),  
 n=n(),   
 pct=mg/n,  
 pct\_close=mgc/n,  
 mean\_pct\_change=mean(pct\_change),  
 mean\_abs\_error=mean(abs\_error),  
 mean\_log\_abs\_error=mean(log\_abs\_error)) |>   
 mutate(accuracy\_level = factor(  
 case\_when(  
 mean\_abs\_error < 0.02 ~ "Exact match",  
 mean\_abs\_error <= 0.05 ~ ".01-5% error",  
 TRUE ~ "Over 5% error" # Capture all remaining cases  
 ),   
 levels = c("Exact match", "01-5% error", "Over 5% error"),  
 ordered = TRUE  
 ))

##| label: tbl-s2-agg  
##| tbl-cap: "Study 2: Summary of planning accuracy by reference class. The table shows performance as both the % of trials where participants matched the goal, and the mean absolute error from the target reduction goal."  
  
  
# overall pct of subjects who matched their goal  
s2\_agg4 |> group\_by('Reference Class' = refClass) |>  
 summarise(  
 #'Avg. % Change' = mean(mean\_pct\_change),  
 '% meeting goal (exact)' = mean(pct),  
 '% meeting goal (close match)' = mean(pct\_close),  
 'Abs. Deviation' = median(mean\_abs\_error),  
 'Log Abs. Deviation' = (median(mean\_log\_abs\_error)),  
) |> mutate(across(where(is.numeric), \(x) round(x, 3))) %>%   
 kable(escape=FALSE,booktabs=TRUE,align=c("l"))

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 5: Experiment 2: Summary of planning accuracy by reference class. The table shows performance as both the % of trials where participants matched the goal, and the mean absolute error from the target reduction goal   | Reference Class | % meeting goal (exact) | % meeting goal (close match) | Abs. Deviation | Log Abs. Deviation | | --- | --- | --- | --- | --- | | kWh | 0.44 | 0.52 | 0.02 | -3.9 | | Percentage | 0.28 | 0.42 | 0.06 | -3.2 | | USD | 0.20 | 0.29 | 0.10 | -2.4 | |

s2\_ld <- ggplot(s2\_agg, aes(y = refClass, x = log\_abs\_error, fill = refClass)) +  
 geom\_density\_ridges(aes(col = refClass), alpha = 0.2, scale = 0.5,  
 jittered\_points = TRUE, point\_alpha = 0.7,point\_size=.4,  
 position = position\_raincloud(width = 0.05, height = 0.1,  
 ygap = 0.05)) +  
 geom\_boxploth(width = 0.1, alpha = 0.3, outlier.shape = NA, show.legend = FALSE) +  
 #scale\_y\_discrete(expand = expansion(mult = c(0.2, 0.4))) +  
 # guides(fill = "none", color = guide\_legend(reverse = TRUE)) +  
 guides(fill = "none") +  
 labs(x = "Log Absolute Deviation", y = "Reference Class", color = "Reference Class") +  
 theme(legend.position = "top")  
  
  
s2\_ldc <- ggplot(s2\_agg, aes(y = refClass, x = log\_abs\_error, fill = calc)) +  
 geom\_density\_ridges(aes(col = calc), alpha = 0.2, scale = 0.5,  
 jittered\_points = TRUE, point\_alpha = 0.7,point\_size=.4,  
 position = position\_raincloud(width = 0.05, height = 0.1,  
 ygap = 0.05)) +  
 geom\_boxploth(width = 0.1, alpha = 0.3, outlier.shape = NA, show.legend = FALSE) +  
 scale\_color\_brewer(palette = "Set1") +  
 scale\_fill\_brewer(palette = "Set1") +  
 guides(fill = "none") +  
 labs(x = "Log Absolute Deviation", y = "Reference Class", color = "") +  
 theme(legend.position = "top")  
  
s2\_ld + s2\_ldc

|  |
| --- |
| Figure 5: Experiment 2: Distribution of the log of the absolute error between the participant’s action plan and the reduction goal across different reference class conditions (kWh, Percentage, USD). The right side plots are further separated by calculator usage. A lower log absolute error suggests higher planning accuracy. |

|  |
| --- |
| Table 6  ##| tbl-cap: "Study 2: Ordinal Regression Model Results."  ordinal\_model\_s2\_logit <- brm(  accuracy\_level ~ refClass + calc+pct\_goal+rounded + (1|id)+ (1|state),  data = s2\_agg,  family = cumulative("logit"),  cores = 4,  iter = 3000,  control = list(adapt\_delta = 0.99), # Recommended for ordinal models  prior = c(prior(normal(0, 2), class = "Intercept"), # Priors for thresholds  prior(normal(0, 2), class = "b")), # Priors for predictors  file = paste0(here::here("data/model\_cache",'s2\_acc3\_add.rds')) # Cache for efficiency ) #summary(ordinal\_model\_s2\_logit)   t2 <- as.data.frame(describe\_posterior(ordinal\_model\_s2\_logit, centrality = "Mean"))[, c(1,2,4,5,6)] |>   setNames(c("Parameter", "Estimate", "CI\_Lower", "CI\_Upper", "pd")) |>   mutate(Parameter = stringr::str\_remove(Parameter, "b\_")) |>   kable(escape = FALSE, booktabs = TRUE, align = c("l"), row.names = FALSE)   # pred\_summary\_s2 <- ordinal\_model\_s2\_logit %>% # epred\_draws(newdata = s2\_agg, re\_formula = NA,ndraws=200) %>% # # group\_by("Reference Class"=refClass, rounded, "% Goal"=pct\_goal, Category=.category) %>% # group\_by("Reference Class"=refClass, Category=.category) %>% # summarise( # mean\_prob = mean(.epred), # lower\_ci = quantile(.epred, 0.025), # upper\_ci = quantile(.epred, 0.975), # .groups = "drop" # )  #pred\_summary\_s2 |> kable(escape=FALSE,booktabs=TRUE,align=c("l"), row.names = FALSE)  or2 <- as.data.frame(fixef(ordinal\_model\_s2\_logit)[,-2])|> as.data.frame() %>%  rownames\_to\_column(var = "Parameter") %>%  mutate(across(where(is.numeric), exp)) |>  filter(!stringr::str\_detect(Parameter, "Intercept")) |>   # rename columns to |comparison | odds\_ratio| ci\_lower| ci\_upper|  rename(comparison = Parameter, odds\_ratio = Estimate, ci\_lower = Q2.5, ci\_upper = Q97.5) |>  # rename levels to Percentage vs kWh; USD vs kWh; Rounded vs Not; 15% Goal vs 10% Goal  mutate(comparison = case\_when(  str\_detect(comparison, "refClassPercentage") ~ "Percentage vs kWh",  str\_detect(comparison, "refClassUSD") ~ "USD vs kWh",  str\_detect(comparison, "roundedRounded") ~ "Rounded vs Not",  str\_detect(comparison, "pct\_goal15%") ~ "15% Goal vs 10% Goal",  TRUE ~ comparison  )) |> kable(escape=FALSE,booktabs=TRUE,align=c("l")) |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 7: **Experiment 2.** Parameter estimates from the ordinal regression model. Positive coefficients for refClass predictors indicate increased likelihood of falling into higher error categories relative to the kWh baseline.   | Parameter | Estimate | CI\_Lower | CI\_Upper | pd | | --- | --- | --- | --- | --- | | Intercept[1] | -1.45 | -2.85 | -0.07 | 0.98 | | Intercept[2] | 1.26 | -0.09 | 2.65 | 0.97 | | refClassPercentage | 1.02 | -0.63 | 2.71 | 0.89 | | refClassUSD | 2.27 | 0.53 | 3.98 | 0.99 | | calcNoCalculator | 4.10 | 2.20 | 6.06 | 1.00 | | pct\_goal15% | -0.39 | -0.81 | 0.04 | 0.96 | | roundedRounded | -0.53 | -0.96 | -0.11 | 0.99 | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 8: **Experiment 2.** Odds ratios for group comparisons. Odds ratios greater than 1 indicate increased odds of falling into a worse accuracy category compared to the comparison condition.   | comparison | odds\_ratio | ci\_lower | ci\_upper | | --- | --- | --- | --- | | Percentage vs kWh | 2.78 | 0.53 | 15.0 | | USD vs kWh | 9.68 | 1.69 | 53.4 | | calcNoCalculator | 60.37 | 9.02 | 426.4 | | 15% Goal vs 10% Goal | 0.68 | 0.44 | 1.0 | | Rounded vs Not | 0.59 | 0.38 | 0.9 | |

As in Experiment 1, accuracy was categorized into three ordinal levels: “Exact match” (0% error), “0.01-5% error,” and “Over 5% error”. The analyses for Experiment 2 employed a Bayesian ordinal regression model to examine the probability of falling into one of three accuracy categories (exact match, minor deviations, or substantial deviations) as a function of the reference class condition (kWh, Percentage, USD), while including pct\_goal (10% vs. 15%), rounded (exact vs. rounded usage data), and calculator usage as additional predictors. Random intercepts were specified for both participant and state,

The ordinal regression analysis revealed that the USD reference class significantly increased the odds of higher error categories compared to the kWh reference class (OR = 9.68, 95% CI: [1.69, 53.4]). Participants in the USD condition were therefore substantially more likely to deviate from the target energy reduction goal compared to those in the kWh condition. In contrast, the Percentage condition’s odds ratio relative to kWh was more uncertain (OR = 2.78, 95% CI: 0.53, 15.0), indicating that although there may be a trend toward reduced accuracy in the Percentage condition, the evidence was not definitive.

We also found that using rounded numbers modestly improved accuracy (b = -0.53, 95% CI: [-0.96, -0.11]), with participants having 0.59 times the odds of falling into a worse accuracy category when working with rounded values. The more challenging 15% reduction goal was associated with slightly better performance compared to the 10% goal (b = -0.39, 95% CI: [-0.81, 0.04]), though this effect was relatively small. Consistent with Experiment 1, the use of a calculator had a large and significant effect on accuracy. The coefficient for calcNoCalculator was 4.10 (95% CI: 2.20, 6.06), and the corresponding odds ratio was 60.37 (95% CI: 9.02, 426.4), indicating that participants who did not use a calculator were substantially more likely to fall into higher error categories.

[Figure 6](#fig-s2-ame) shows the marginal effects of refClass on each level of accuracy\_level. These results reveal that switching from kWh to Percentage decreased the probability of an “Exact match” by an average of 7.0 percentage points (95% CI: -19.2, 4.2) and increased the probability of “Over 5% error” by 6.9 percentage points (95% CI: -4.5, 18.6). Similarly, switching from kWh to USD decreased the probability of an “Exact match” by 15 percentage points (95% CI: -26.7, -3.3) and increased the probability of “Over 5% error” by 16.5 percentage points (95% CI: 3.7, 29.3).

library(ggtext)  
  
  
set.seed(133)  
ame2 <- avg\_slopes(  
 ordinal\_model\_s2\_logit,   
 variables = "refClass",ndraws=850  
)  
  
# Add annotations to the data frame  
ame2\_annotated <- ame2 %>%  
 mutate(label = sprintf("%.1f%%", estimate \* 100))  
  
ggplot(ame2\_annotated, aes(x = estimate, y = group, color = contrast, group = contrast)) +  
 geom\_point(size = 3, alpha=.6,position = position\_dodge(width = 0.5)) +  
 geom\_errorbarh(aes(xmin = conf.low, xmax = conf.high), height = 0.2, alpha=.5,  
 position = position\_dodge(width = 0.5)) +  
 geom\_vline(xintercept = 0, linetype = "dashed",alpha=.5) +  
 labs(  
 x = "Average Marginal Effect",  
 y = "Accuracy Level",  
 color = "Comparison",  
 title = "Average Marginal Effects of refClass on Accuracy Levels"  
 ) +  
 # Add annotations  
 geom\_text(  
 aes(label = label),  
 color = "black", size = 3.5, hjust = -0.3, vjust = -0.5,  
 position = position\_dodge(width = 0.5)  
 ) +  
 theme\_minimal()

|  |
| --- |
| Figure 6: Experiment 2. Average marginal effects of reference class on accuracy levels (Experiment 2). The points represent the average change in the probability of each accuracy level when switching from the kWh reference class to Percentage (red) or USD (green). Error bars indicate 95% credible intervals. The results show that, compared to kWh, the Percentage format decreases the probability of an “Exact match” by 7.0% and increases the probability of “Over 5% error” by 6.9%. The USD format has a larger negative effect on “Exact match” (-14.7%) and a larger positive effect on “Over 5% error” (+16.5%). The effects on the “0.01-5% error” category are near zero for both comparisons. |

pp\_check(ordinal\_model\_s2\_logit, type = "bars\_grouped", group="refClass", fatten = 2,ndraws=400) +  
 scale\_x\_continuous("Response Category", breaks = 1:3,   
 labels = c("Exact", "0.01-5%", ">5%")) +  
 scale\_y\_continuous(expand = expansion(mult = c(0, 0.05))) +  
 ggtitle("Posterior Predictive Check by Reference Class") +  
 theme\_minimal() +  
 #scale\_fill\_manual(values = c("kWh" = "#66c2a5", "Percentage" = "#fc8d62", "USD" = "#8da0cb"),name="Reference Class") +  
 theme(  
 legend.background = element\_blank(),  
 legend.position = "bottom",  
 panel.grid.minor = element\_blank(),  
 axis.text.x = element\_text(angle = 45, hjust = 1))

|  |
| --- |
| Figure 7: Experiment 2. Posterior predictive check of the bayesian regression model, faceted by reference class. The bars represent the observed frequencies of each accuracy level within each reference class. The points represent the model’s predicted proportions, with error bars indicating 95% credible intervals. |

s2\_els\_log\_error <- brm(  
 log\_abs\_error ~ els + (1|id) + (1|state),  
 data = s2\_agg,  
 family = gaussian(),  
 cores = 4,  
 iter = 3000,  
 control = list(adapt\_delta = 0.97),   
 prior = c(prior(normal(0, 3), class = "Intercept"),   
 prior(normal(0, 3), class = "b")),   
 file = paste0(here::here("data/model\_cache",'s2\_els\_log\_error.rds'))   
)  
#summary(s2\_els\_log\_error)  
# Regression Coefficients:  
# Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
# Intercept -1.10 0.27 -1.62 -0.56 1.00 975 1675  
# els -3.21 0.35 -3.89 -2.52 1.00 837 1508  
  
  
conditional\_effects\_plot <- conditional\_effects(s2\_els\_log\_error)  
plot\_data <- conditional\_effects\_plot[[1]]  
  
ggplot(plot\_data, aes(x = els, y = estimate\_\_)) +  
 geom\_line(color = "blue", size = 1) +  
 geom\_ribbon(aes(ymin = lower\_\_, ymax = upper\_\_), alpha = 0.2) +  
 labs(  
 x = "Energy Literacy Score",  
 y = "Log Absolute Error",  
 title = "Conditional Effect of Energy Literacy on Log Absolute Error"  
 ) +  
 theme\_minimal()

|  |
| --- |
| Figure 8: Experiment 2. Conditional effect of energy literacy on log absolute error. The plot shows the relationship between energy literacy score and log absolute error, controlling for random effects of participant and state. Higher energy literacy scores are associated with smaller deviations from the target reduction goal, indicating more accurate planning. |

We once again examined the effect of energy literacy on planning accuracy. A Bayesian linear regression model was fit with log-transformed absolute error as the outcome variable and energy literacy score as the predictor, controlling for random effects of participant and state: log\_abs\_error ~ els + (1|id) + (1|state). This revealed a significant negative relationship between energy literacy and log absolute error (Estimate = -3.21, 95% CI: -3.89 to -2.52), indicating that participants with higher energy literacy scores tended to have smaller deviations from the target reduction goal, and thus more accurate plans overall ([Figure 8](#fig-s2-els)).

## Experiment 2: Discussion

Experiment 2 aimed to build upon the findings of Experiment 1, not only by replicating the core manipulation of reference class, but also by incorporating additional variables that might influence planning accuracy. These included goal difficulty and the way that the prior year’s usage was presented (rounded or exact). The results largely converged with those of Experiment 1, providing further converging evidence that presenting energy reduction goals in absolute units (kWh) facilitates more accurate planning compared to percentage-based or monetary formats.

Taken together, the results of Experiment 2 provide further support for the hypothesis that presenting energy reduction goals in absolute units (kWh) leads to more accurate planning compared to percentage-based or monetary formats.

The finding that the more challenging 15% reduction goal was associated with a slight improvement in accuracy is counterintuitive. It may be that participants put more effort into the task under this condition, or perhaps this is an artifact of the way that the task was presented. However, this effect was relatively small and thus should be explored in future research to better understand its underlying mechanisms. Furthermore, the magnitude of the effect size of this manipulation should be examined to better understand the practical implications of goal difficulty for energy conservation.

The large and significant effect of calculator use, consistent across both experiments, underscores the crucial role of tools that individuals are likely to employ in real-world settings. Finally, the consistent relationship between energy literacy and accuracy, observed across both experiments, highlights the potential value of educational interventions aimed at improving consumers’ understanding of energy concepts.

# General Discussion

This study examined how different numerical representations of energy reduction goals influence consumers’ ability to create accurate energy conservation plans. Across two experiments, the findings consistently demonstrated that presenting reduction goals in absolute units (kWh) significantly enhanced planning accuracy compared to percentage-based or monetary formats. Furthermore, participants with higher energy literacy exhibited more precise planning across all conditions, underscoring the critical role of domain-specific knowledge in shaping decision-making outcomes.

These results align with the broader literature on numerical cognition, which suggests that the format in which information is presented can profoundly affect comprehension and decision-making (Gigerenzer & Edwards, 2003; Reimer et al., 2015). However, our study extends beyond simple estimations or judgments, demonstrating that the advantages of absolute units persist even in a more complex, multi-step planning task. Moreover, the consistent superiority of kWh over both percentages and USD provides novel insights into the specific challenges of energy-related decision-making. While prior research has suggested that consumers may prefer monetary feedback (Karjalainen, 2011; Nemati & Penn, 2020) or that monetary framing can improve long-term appliance choices (Blasch et al., 2019), our findings indicate that when it comes to allocating specific usage reductions across appliances, absolute units are most effective. This may be because kWh provide a more direct and less ambiguous representation of energy quantities, facilitating the necessary calculations for accurate planning. Nevertheless, it is unclear whether the observed benefits of absolute units (i.e., kWh) are genuinely attributable to their absolute nature, or if other inherent characteristics of these units might be driving the effects. Furthermore, the percentage-based reduction targets, while potentially more salient from a goal-setting perspective, also resulted in poorer planning outcomes compared to kWh. Such a finding suggests that reliance on percentages can further complicate calculations by adding unnecessary transformations in the problem-solving process, in line with the “off by 100% bias” found in Fisher & Mormann (2022), where they show that individuals often misunderstand percentage changes greater than 100%. The relative ease in using absolute values to arrive at an accurate plan suggests that these representations of the planning task result in the most accurate planning strategies because the number of steps required to perform the required calculations are simpler.

Moreover, the consistent positive relationship observed between energy literacy and planning accuracy across both experiments underscores the importance of baseline knowledge in effectively navigating energy-related information. Individuals with higher levels of energy literacy demonstrated a greater capacity to formulate accurate conservation plans, irrespective of the information format presented. This finding aligns with previous research highlighting the role of numeracy and domain-specific knowledge in improving judgments related to energy consumption (Attari et al., 2010). It further suggests that interventions aimed at enhancing consumers’ fundamental understanding of energy concepts could yield significant benefits in improving the effectiveness of energy conservation efforts.

## Limitations

* Each participant only completed 4 action plans
* calculator use not controlled.

# References

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, *25*(3), 273–291. <https://doi.org/10.1016/j.jenvp.2005.08.002>

Attari, S. Z., DeKay, M. L., Davidson, C. I., & Bruine De Bruin, W. (2010). Public perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences*, *107*(37), 16054–16059. <https://doi.org/10.1073/pnas.1001509107>

Bednar, D. J., & Reames, T. G. (2020). Recognition of and response to energy poverty in the United States. *Nature Energy*, *5*(6), 432–439. <https://doi.org/10.1038/s41560-020-0582-0>

Blasch, J., Filippini, M., & Kumar, N. (2019). Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resource and Energy Economics*, *56*, 39–58. <https://doi.org/10.1016/j.reseneeco.2017.06.001>

Bürkner, P.-C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1–28. <https://doi.org/10.18637/jss.v080.i01>

Canfield, C., Bruine De Bruin, W., & Wong-Parodi, G. (2017). Perceptions of electricity-use communications: Effects of information, format, and individual differences. *Journal of Risk Research*, *20*(9), 1132–1153. <https://doi.org/10.1080/13669877.2015.1121909>

DeWaters, J. E., & Powers, S. E. (2011). Energy literacy of secondary students in New York State (USA): A measure of knowledge, affect, and behavior. *Energy Policy*, *39*(3), 1699–1710. <https://doi.org/10.1016/j.enpol.2010.12.049>

Farghali, M., Osman, A. I., Mohamed, I. M. A., Chen, Z., Chen, L., Ihara, I., Yap, P.-S., & Rooney, D. W. (2023). Strategies to save energy in the context of the energy crisis: A review. *Environmental Chemistry Letters*, *21*(4), 2003–2039. <https://doi.org/10.1007/s10311-023-01591-5>

Fischer, C. (2008). Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency*, *1*(1), 79–104. <https://doi.org/10.1007/s12053-008-9009-7>

Fisher, M., & Mormann, M. (2022). The Off by 100% Bias: The Effects of Percentage Changes Greater than 100% on Magnitude Judgments and Consumer Choice. *Journal of Consumer Research*, *49*(4), 561–573. <https://doi.org/10.1093/jcr/ucac006>

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *BMJ*, *327*(7417), 741–744. <https://doi.org/10.1136/bmj.327.7417.741>

Gill, C. A., Atlas, S. A., Hardisty, D. J., & Scott, S. P. (2022). Consumer matching costs to context: Status quo bias, temporal framing, and household energy decisions. *Journal of Consumer Behaviour*, *21*(5), 1018–1027. <https://doi.org/10.1002/cb.2051>

Karjalainen, S. (2011). Consumer preferences for feedback on household electricity consumption. *Energy and Buildings*, *43*(2-3), 458–467. <https://doi.org/10.1016/j.enbuild.2010.10.010>

Memmott, T., Carley, S., Graff, M., & Konisky, D. M. (2021). Sociodemographic disparities in energy insecurity among low-income households before and during the COVID-19 pandemic. *Nature Energy*, *6*(2), 186–193. <https://doi.org/10.1038/s41560-020-00763-9>

Nemati, M., & Penn, J. (2020). The impact of information-based interventions on conservation behavior: A meta-analysis. *Resource and Energy Economics*, *62*, 101201. <https://doi.org/10.1016/j.reseneeco.2020.101201>

Reimer, T., Jones, C., & Skubisz, C. (2015). Numeric Communication of Risk. In *The SAGE handbook of risk communication* (pp. 167–179).

Team, R. C. (2020). *R: A Language and Environment for Statistical Computing*. R: A Language and Environment for Statistical Computing.

Tonke, S. (2024). Providing procedural knowledge: A field experiment to encourage resource conservation in Namibia. *Journal of Development Economics*, *166*, 103202. <https://doi.org/10.1016/j.jdeveco.2023.103202>

Weber, P., Binder, K., & Krauss, S. (2018). Why Can Only 24% Solve Bayesian Reasoning Problems in Natural Frequencies: Frequency Phobia in Spite of Probability Blindness. *Frontiers in Psychology*, *9*, 1833. <https://doi.org/10.3389/fpsyg.2018.01833>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>