nature portfolio

https://doi.org/10.1038/s41586-021-03380-y

Supplementary information

People systematically overlook subtractive changes

In the format provided by the authors and unedited

Supplementary Information

for

People Systematically Overlook Subtractive Changes

Gabrielle S. Adams, Benjamin A. Converse, Andrew H. Hales, and Leidy Klotz

Table of Contents

Supplementary Information for *People Systematically Overlook Subtractive Changes*

n.b. All data, code, and materials are available at:
https://osf.io/7v6r2/?view_only=6bee983832354e9c99daf58590b9d43c

1. Participant Exclusions, Analyses, and Results for Experiments 1—8

1.1 Experiment 1 exclusions, analyses, and results

We recruited 203 participants from people passing by a table in a highly-trafficked area of a large public university in the United States. Six sessions were excluded for procedural or random-assignment protocol violations, yielding a final sample of $N = 197$ participants. We preregistered these exclusions and an analysis plan at https://osf.io/rkqvw/.

The dependent measure of interest was whether each participant produced a transformed structure that had fewer total pieces than in the original. As reported in the main text, the subtraction cue significantly increased the percentage of participants whose transformation had fewer pieces than the original. In the control condition, 41% produced a subtractive transformation and in the subtraction-cue condition 61% produced a subtractive transformation, $\chi^2(1, N = 197) = 7.72$, $p = .005$, $\varphi = 0.20$. (Results are similar if the six cases that violated experimental protocol are included, $\chi^2(1, N = 203) = 7.45$, $p = .006$, $\varphi = 0.19$.)

Note that the dependent measure in the pre-registered analysis was “whether the participant removed any pieces (removed v. did not remove).” With this measure, any transformed tower that had even a single piece removed is counted as a subtractive transformation. This pre-registered analysis also supports the general conclusion that the cue increases subtraction: In the control condition, 48% removed at least one piece while transforming the tower, whereas 63% did so in the subtraction-cue condition, $\chi^2(1, N = 197) = 4.29$, $p = .038$, $\varphi = 0.15$. (Results are similar if we include the six cases that violated protocol, $\chi^2(1, N = 203) = 4.04$, $p = .045$, $\varphi = 0.14$.)

While analysis using either dependent measure supports the same conclusion, we determined that the dependent measure focused on the net total of blocks in the transformed structure is more appropriate because it allows that a participant might be focused on adding (i.e., using a mental model of additive change), but still remove a piece in the service of this focus on adding. For example, a transformed structure with an entire added layer but a single "leg" removed to make room for this layer is most accurately categorized as an additive transformation.

1.2 Experiment 2 exclusions, analyses and results

We recruited 262 participants on Amazon Mechanical Turk (MTurk). Among the respondents, 96 accessed the survey but withdrew before consenting (following notification that the study included a writing component), 14 provided no valid responses, and 5 were from duplicated IP addresses, yielding a final sample of $N = 147$ participants.

The dependent measure of interest was whether each participant provided at least one subtractive idea to improve the miniature golf hole. A higher percentage of participants listed at least one subtractive change idea in the cue condition (44%) compared to the no-cue condition (30%), but not significantly so, $\chi^2(1, N = 147) = 2.97$, $p = .085$, $\varphi = 0.14$. We did not find evidence that the cue manipulation affected rates of additive change ideas, with no significant difference between the percentage who provided at least one additive change idea in the no-cue condition (95%) compared to the cue condition (93%), $\chi^2(1, N = 147) = .33$, $p = .564$, $\varphi = 0.05$. Given the suggestive but inconclusive result for subtractive change ideas, a close replication was conducted in experiment 3.

Checks supported the assumption that most participants noticed the cue manipulation. In the control condition, 59% correctly reported there being no cue, with another 34% reporting that they did not recall, and the remainder (6%) incorrectly reporting there had been a cue. In the cue condition, 76% correctly reported seeing a cue, 13% did not recall, and the remainder (10%) incorrectly reported there had not been a cue. An exploratory (*post hoc*) analysis including only participants who passed the manipulation check (i.e., those in the control condition who indicated "no" to seeing a cue, and those in the cue condition who said "yes"), suggested a significant effect of cue on rates of subtraction: $\chi^2(1, N = 99) = 4.05$, $p = .044$, $\varphi = 0.20$.

1.3 Experiment 3 exclusions, analyses, and results

We recruited 336 participants on MTurk. Among the respondents, 27 accessed the survey but withdrew before consenting (following notification that the study included a writing component), 127 provided zero valid responses, 4 were from duplicated IP addresses, and 13 failed an attention check, yielding a final sample of N =165.

The dependent measure of interest was whether participants provided at least one subtractive change idea to improve the miniature golf hole. Significantly more participants offered at least one subtractive change in the cue condition (42%) than in the no-cue condition (26%), $\chi^2(1, N = 165) = 4.62$, $p = .032$, $\varphi = 0.17$. As in experiment 2, the cue did not change percentages of participants who provided at least one additive change idea; no-cue condition (80%) and cue condition (81%), $\chi^2(1, N = 165) < .01$, $p = .960$, $\varphi < .01$.

We combined results from experiments 2 and 3 into a meta-analysis (assuming random effects[1]), to assess the robustness of the cue manipulation in eliciting subtractive change ideas. Pooling the effect size from both experiments revealed a reliable effect whereby the cue increases one's likelihood of producing a subtractive change idea, OR = 1.93, 95% CI [1.20, 3.08], z = 2.73, p = .006; but there was not evidence that the cue affected one's likelihood of producing an additive change idea, OR = 0.92, 95% CI [0.47, 1.80], z = -0.24, p = .810.

1.4 Experiment 4 exclusions, analyses, and results

We recruited 658 participants on MTurk. Among the respondents, 80 accessed the survey but withdrew before consenting (following an early instruction screen notifying participants that the study included a writing component), 179 provided zero valid responses or withdrew during the task, 3 were from duplicated IP addresses, and 27 failed an attention check (the same one used in experiment 3), yielding a final sample of $N = 369$.

The dependent variable of interest was whether participants provided at least one subtractive change idea to improve the miniature golf hole. A logistic regression predicted the likelihood of subtraction from the cue condition (cue/no cue), the goal condition (make-better/make-worse), and their interaction. The regression revealed a main effect of cue, $b = 1.09$, $z = 4.75$, $p < .001$, but no effect of goal, $b = .22$, $z = .97$, $p = .332$, and no interaction, $b = -.31$, $z = -.68$, $p = .495$. The cue increased the likelihood of producing a subtractive change idea both when the goal was to make the hole better (no-cue = 21%, cue = 48%, $\chi^2(1, N = 171) = 13.63$, $p < .001$, $\varphi = 0.28$), and when the goal was to make the hole worse (no-cue = 28%, cue = 50%, $\chi^2(1, N = 198) = 9.71$, $p = .002$, $\varphi = 0.22$).

A logistic regression also modeled the likelihood of addition as a function of the cue condition (cue/no cue), the goal condition (make-better/make-worse), and their interaction. The regression revealed no effects of either the cue ($b = .31$, $z = 1.04$, $p = .300$), the goal condition ($b = -.41$, $Z = -1.36$, $p = .174$), or their interaction ($b = .99$, $z = 1.63$, $p = .103$).

1.5 Standardized Median Rank Difference analysis of experiments 2—4

Our focal analytic approach in experiments 2—4 accounts for the presence or absence of subtractive ideas but does not address the order in which those ideas emerged. Given that 50% or more of the sample in each condition from those experiments did not report any subtractive ideas, this straightforward categorical analysis seems like an appropriate way to describe and analyze the results. To investigate order effects in more detail, however, research on the role of query order in value construction provides a potential analytic strategy, the Standardized Median Rank Difference (SMRD)[2].

The SMRD approach compares the median rank of two different categories of ideas. To accommodate this, we categorized each response as either subtractive or non-subtractive (the latter collapsed the additions and changes categories). Applied to our data, the SMRD analysis uses the formula: $2(MR_a - MR_s)/n$, where MR_a = median rank of non-subtractive ideas, MR_s = median rank of subtractive ideas, and n = the number of ideas a participant generated. This formula produces scores ranging from a minimum of -1 (where all non-subtractive ideas come before any subtractive idea) to a maximum of 1 (where all subtractive ideas come before any non-subtractive idea). Intuitively, then, higher numbers represent higher relative accessibility of subtractive than non-subtractive ideas. For example, a participant who generated the sequence *Add, Add, Subtract, Add* would have a score of $2*(2 - 3)/4 = -0.50$, whereas a participant who generated the sequence *Subtract, Subtract, Subtract, Add* would have a score of $2*(4 - 2)/4 = 1$.

Note that without an adjustment, the SMRD does not accommodate any list that lacks one of the categories. The formula would be undefined for a participant who provided a list with only additive ideas (because no median rank can be determined for subtractive ideas) or vice versa. Following the original approach by Johnson and colleagues, we applied a correction that results in such cases receiving the maximum possible score for people who only subtracted (SMRD = 1), and the minimum possible score for people who only added/changed (SMRD = -1). We note that this is a very strong assumption. Its effect is to assume that the next idea in a participant's list would have been from the other category. So, for example, a participant whose list was *Add, Add, Add* would be treated in the analysis as if his list was *Add, Add, Add, Subtract*. One consequence is that a participant who provided three additive ideas would receive the same score (SMRD = -1) as a participant who provided three additive ideas followed by three subtractive ideas (*SMRD* also = -1). This makes it an extremely conservative test for our hypothesis, and a good robustness check for our featured analytic approach.

The distribution of SMRD scores was highly non-normal (with many -1 scores, reflecting the high percentage of participants who did not report any subtractions). We therefore report *t*-tests and nonparametric tests using Wilcoxon rank sums. In experiment 2, subtractive ideas appeared descriptively earlier in the cue condition ($M = -.53$, $SD = .72$) than in the no-cue condition ($M = -.69$, $SD = .65$), but not to a statistically significant degree, $t(135.96) = -1.44$, $p = .152$, $d = -.24$, $W = 2{,}371.5$, $p = .119$. Similarly, in experiment 3, subtractive ideas appeared descriptively but not significantly earlier in the cue condition ($M = -.53$, $SD = .77$) than in the no-cue condition ($M = -.69$, $SD = .66$), $t(152.65) = -1.44$, $p = .151$, $d = -.23$, $W = 3{,}069$, $p = .171$. In the improvement-goal condition of Experiment 4, subtractive ideas appeared significantly earlier in the cue condition ($M = -.40$, $SD = .82$) than in the no-cue condition ($M = -.81$, $SD = .49$), $t(149.98) = -4.00$, $p < .001$, $d = -.60$, $W = 2{,}677.5$, $p < .001$; and in the make-it-worse condition of Experiment 4, subtractive ideas appeared significantly earlier in the cue condition ($M = -.43$, $SD = .74$) than in the no-cue condition ($M = -.68$, $SD = .67$), $t(195.73) = -2.47$, $p = .014$, $d = -.35$, $W = 3{,}828$, $p = .002$. A meta-analysis of all four studies (with the *improve* and *make-worse* conditions of Experiment 4 entered as separate studies) showed a clear tendency for the cue to cause subtractive ideas to be generated sooner in the sequence than in a no-cue condition; specifically, the pooled average SMRD score was -0.72 in the no-cue condition and -0.47 in the cue condition, Hedges $g = -.35$, $Z = -4.30$, $p < .001$. Based on these analyses, it appears that on average people do tend to think of subtractive ideas relatively sooner when they are prompted with a cue than they otherwise would.

1.6 Experiment 5 exclusions, analyses, and results

We recruited 317 participants from MTurk. Among these respondents, 16 did not complete the task, and 2 had duplicated IP addresses, yielding a final sample of $N = 299$.

The dependent measure of interest was the likelihood that a participant would provide the correct transformation, which required subtracting filled in blocks from the grid. As predicted, participants were significantly less likely to provide the correct transformation in the control condition (49%), than in the repeated-search condition (63%), $\chi^2(1, N = 299) = 5.87$, $p = .015$, $\varphi = 0.15$.

1.7 Experiment 6 exclusions, analyses, and results

We recruited 313 participants from MTurk. Among these respondents, 11 did not complete the task and 3 had duplicated IP addresses, yielding a final sample of $N = 299$.

The dependent measure of interest was the number of trials (from zero to four) on which each participant provided the correct transformation, which required subtracting filled-in blocks from the grid. Participants in the higher-cognitive load condition provided significantly fewer subtractive transformations ($M = 2.20$, $SD = 1.81$) than did participants in the lower-cognitive load condition ($M = 2.64$, $SD = 1.70$), $t(294.96) = 2.19$, $p = .029$, $d = .25$. Because the data were bimodally distributed (many participants either added on all four trials or subtracted on all four trials), a non-parametric Wilcoxon rank sum test was performed, which also indicated a significant difference between conditions, $W = 12{,}710$, $p = .027$.

In an additional analysis, we examined whether conditions differed in the likelihood of producing at least one subtraction across the four grids. The percentage who did so was significantly greater in the lower-load condition (75%) than in the higher-load condition (64%), $\chi^2(1, N = 299) = 4.50$, $p = .034$, $\varphi = 0.12$.

<u>1.8 Experiment 7 exclusions, analyses, and results</u>

We recruited 286 undergraduate students enrolled in introductory psychology courses at a large public university. No participant responses were excluded from analysis.

The dependent measure of interest was the number of trials (from zero to four) on which each participant provided the correct transformation, which required subtracting filled in blocks from the grid. As predicted, participants provided the correct transformation significantly less often in the higher-load condition ($M = 2.49$, $SD = 1.75$) than in the lower-load condition ($M = 2.94$, $SD = 1.53$), $t(281.73) = 2.32$, $p = .021$, $d = .27$. Because the data were bimodally distributed (many participants either added on all four trials or subtracted on all four trials), we also conducted a non-parametric Wilcoxon rank sum test, which also indicated a significant difference between conditions, $W = 11{,}556$, $p = .035$.

Responses to the process check support the assumption that participants experienced the task as more challenging in the higher-load condition ($M = 4.46$, $SD = 1.28$) than in the lower-load condition ($M = 1.57$, $SD = .91$), $t(262.06) = -22.13$, $p < .001$, $d = -2.60$.

We also examined whether conditions differed on the likelihood of producing at least one subtraction across the four grids. The percentage who did so was significantly greater in the lower-load condition (84%) than in the higher-load condition (73%), $\chi^2(1, N = 286) = 5.75$, $p = .017$, $\varphi = 0.14$.

1.9 Experiment 8 exclusions, analyses and results

We recruited 600 undergraduate students from a large public university in the United States. Among these respondents, 17 were excluded for duplicate responding, 12 for not responding to all four grid tasks, and 3 for making no changes to the grids, leaving a final sample of $N = 568$.

The dependent measure of interest was the number of trials (from zero to four) on which each participant provided the correct transformation, which required subtracting filled in blocks from the grid. We did not find a significant difference between the number of subtractive transformations provided by participants in the higher-load condition ($M = 2.57$, $SD = 1.71$) and participants in the lower-load condition ($M = 2.74$, $SD = 1.68$), $t(563.86) = 1.22$, $p = .225$, $d = .10$. Because the data were bimodally distributed (many participants either added on all four trials or subtracted on all four trials), we also conducted a non-parametric Wilcoxon rank sum test, which indicated a nonsignificant difference between conditions, $W = 42{,}988$, $p = .133$.

Responses to the manipulation check supported the assumption that the task was experienced as more challenging in the higher-load condition ($M = 4.91$, $SD = 1.38$) than in the lower-load condition ($M = 1.58$, $SD = 1.12$), $t(532.13) = -31.52$, $p < .001$, $d = -2.66$.

We also examined whether conditions differed on the likelihood of producing at least one subtraction across the four grids. The percentage was not significantly different in the lower-load condition (77%) versus the higher-load condition (73%), $\chi^2(1, N = 568) = 1.13$, $p = .287$, $\varphi = 0.04$.

1.10 Meta-analysis of experiments 6—8 and single-trial experiment

As noted in the main text, meta-analysis of experiments 6—8 ($N = 1153$) indicates that participants failed to identify the correct subtractive transformation for more puzzles in the higher- versus lower-load condition, Hedge's $g = .18$, $z = 2.97$, $p = .003$.

Note that, prior to experiments 7 and 8, we conducted a study using the head-moving cognitive load manipulation from experiment 6 with an online sample ($N = 441$; lower-load condition $n = 224$, higher-load condition $n = 217$). This study was not ultimately included as one of the primary experiments because it used a different operationalization, asking participants to respond to a single grid, rather than to four grids as in experiments 6—8. The dependent measure of interest in this additional study was a dichotomous count of whether participants produced the correct subtractive transformation. This single-trial study did not show an effect of head-moving on rates of subtraction: 44% of participants produced the subtractive solution in the lower-load condition and 42% produced the subtractive solution in the higher-load condition, $\chi^2(1, N = 441) = .15$, $p = .703$, $\varphi = 0.02$. We were concerned about compliance with the cognitive-load instructions, but could not empirically assess this explanation. We therefore moved to an in-lab administration in experiment 7.

Though we could not explain the discrepant finding of the single-trial experiment, we conducted a broader meta-analysis using only participants' first trial from experiments 6—8 and their response in the single-trial study. This meta-analysis similarly supports the overall finding that cognitive load significantly reduced the likelihood of subtracting, total N = 1,594, OR = 0.72, 95% CI [0.58, 0.90], z = -2.85, p = .004.

2. Descriptions of Methods and Results for Studies S1—S12

As described in the introduction of the main text, we conducted a series of observational studies to assess the robustness of our behavioral assumption that subtractive transformations appear less frequently than additive transformations in many contexts. Evidence consistent with this assumption subsequently led to the proposal that people may engage in additive searches as a heuristic default when considering possible changes, which, in turn, motivated experiments 1—8. Results for these motivating observations are summarized in Extended Data Table 1, and each study is described in more detail in the sections below.

Studies S9 and S10 investigated subtracting behavior when stimuli contained more and less extraneous (S9) or anomalous (S10) components. Study S11 was a post-hoc validation study designed to investigate a potential alternative explanation of experiment 1. Study S12 informed a recommendation for future research in the "Discussion" section of the main text, probing the tendency to search for additive versus subtractive changes across samples of participants intentionally selected from three geographic locations: Germany, Japan, and the United States.

2.1 Study S1 – Abstract grid task

In the task for study S1, participants transformed a grid pattern to make it symmetrical. This abstract grid task allowed observation of rates of additive and subtractive transformation in a context with which participants would have no prior experience, and therefore no learned habits or developed preferences specific to the context. This task could be accomplished equally well with an additive transformation (filling squares on one side of the grid), or with a subtractive transformation (unfilling squares on the other side of the grid.)

We recruited 110 participants on MTurk. Of the 110 respondents, 7 did not complete the survey, and 6 were from duplicate IP addresses, yielding a final sample of 97 (39 men, 58 women; M_{age} = 34.19, *SD* = 9.59). Participants received $0.60 for participating.

First, participants viewed a practice grid and read instructions: "Click spaces below to see how squares can be manipulated." Clicking on any square would toggle it between white (blank) and green (filled). After familiarizing themselves with the workspace, participants read, "On the following pages, you will be shown patterns. For each one, your job is to arrange the blocks so that the left side perfectly mirrors the right side, and vice-versa." Participants then saw examples of correct and incorrect transformations. Next, participants continued to the main task, which included six trials. On each trial (presented in random order for each participant), participants saw one grid, with the instruction to: "Arrange the blocks so that the left side perfectly mirrors the right side and vice-versa. ONLY CONTINUE WHEN THE TWO SIDES ARE SYMMETRICAL." After completing all six trials, participants indicated their approach to changing the grids, "*In general,* what was your approach to make the sides symmetrical?", with two response options: "*I added squares until they were symmetrical*" or "*I removed squares until they were symmetrical.*"

The observation of interest in study S1 was a categorical measure of the response that participants produced most frequently across the six trials. We categorized participants' dominant behavior as adding, subtracting, or breaking-even if they did so on four or more out of the six trials. We excluded 3 participants from analysis because they did not have a dominant response (they subtracted in three trials and added in three trials), and 3 other participants because they provided the break-even response (neither added nor subtracted) most often. All other participants provided either an additive or subtractive transformation at least 4 times. Among these 91 remaining participants, 73 participants (80%) added to transform the grids 4 or more times and 18 participants (20%) subtracted to transform the grids 4 or more times. A test of whether 18/91 differs from 50% rejects the null hypothesis that subtractive transformations were as common as additive transformations, two-sided binomial distribution probability, $p < .001$. Findings are similar if data are analyzed based on total number of transformations: overall the 97 participants provided an average of 1.5 subtractive transformations ($SD = 1.8$) out of 6. This average is significantly lower than the average of 3 trials that would be expected if participants were as likely to subtract as to add, one-sample $t(96) = -8.50$, $p < .001$, $d = -.86$. Participants' self-reported responses regarding their approach to changing the grids also support these findings: overall, 75% of participants reported generally making additive changes, and 25% reported generally making subtractive changes, two-sided binomial distribution probability, $p < .001$.

Note that, as an early test of procedural priming effects, we randomly assigned participants to practice on either a grid that began with only white squares ($n = 48$) or a grid that began with only green squares ($n = 49$). There was not evidence that initial experience with the all-white versus all-green initial grid affected the average number of subtractive transformations that participants made, $t(94.94) = .49$, $p = .629$, $d = .10$. Among participants who initially practiced with all-white space, 24% subtracted as their dominant response, and among those who initially practiced with an all-green space, 22% did so, $\chi^2(1, N = 91) = .04$, $p = .848$. Responses to the self-reported question about participants' approach to changing the grids also did not vary by initial experience, $\chi^2(1, N = 97) < .01$, $p = .954$.

2.2 Study S2 – Request for Ideas (“Ours to Shape” — field observation)

The task for study S2 was a naturalistic field observation of suggested changes to a large, public university in the United States. Stakeholders offered suggestions in response to an open call from the incoming university President. This allowed us to observe the prevalence of additive and subtractive action ideas in response to a consequential, real-world stimulus not created by the researchers.

Upon request, university administrators provided us a list of responses that had been submitted on a dedicated website (N = 1,201). To preserve anonymity, the data file we received did not allow us to determine whether ideas came from the same or different person (or IP address). We treated each response as an independent observation but cannot verify this to be the case. Analysis of these archival data was approved by the Institutional Review Board at the University of Virginia.

In the open call for ideas, the President’s office invited university stakeholders (students, faculty, staff, community members, etc.) to offer suggestions in any of three categories: 1) Community (i.e., “How do we strengthen the sense of community [at the university], with the [region], and with our alumni?”, 2) Discovery (i.e., “How do we best enable students to learn about the past, themselves, and the world around them?”), and 3) Service (i.e., “How do we prepare our students for a lifetime of service, regardless of their chosen fields?”). The form allowed submissions up to 700 characters, and respondents could choose to identify their affiliation(s) to the university. (Affiliations reported were: alumnus (48.5%), community member (23.7%), staff (22.4%), student (17.2%), faculty (16.0%), supporter (15.1%) and parent (13.5%). The sum exceeds 100% because respondents could indicate multiple affiliations.)

Two independent coders categorized each suggestion into one of four categories: additive transformations (*the suggestion involved adding elements or resources to the university*), subtractive transformations (*the suggestion involved removing elements or resources from the university*), neither additive nor subtractive (*the suggestions changes or replaces some element of the university without a clear addition or subtraction*), and invalid response (*the suggestion was empty, incomprehensible, or clearly effortless*). Coders were unaware of specific hypotheses or predictions. Coding was reliable (Kappa = .89), and coders resolved disagreements through discussion.

The observation of interest in study S2 was whether suggested ideas were additive or subtractive. We excluded from analysis responses that were categorized as invalid (n = 374). Of the valid responses, 70% were additive, 21% were neither additive nor subtractive, and 8% were subtractive. To directly assess whether additive solutions outnumber subtractive solutions, we removed suggestions that were neither subtractive nor additive (n = 176). Among the remaining 651 suggestions, 70 (11%) were subtractive (e.g., "eliminate legacy admissions," "no rape, sexual abuse, or abuses of power") and 581 (89%) were additive (e.g., "more study abroad grants," "create gender-inclusive housing options"). A test of whether 70/651 differs from 50% rejects the null hypothesis that subtractive ideas were suggested as often as additive ideas, two-sided binomial distribution probability, $p < .001$.

<u>2.3 Study S3 – Block structures</u>

In the task for study S3, participants transformed a small structure made of Lego-like (Duplo) blocks. This allowed observation of additive and subtractive behavior when transforming physical objects.

We recruited participants at a table on a university campus ($N = 42$), with a sign that said, "Do you have 2-5 minutes to stop and play with Legos?" The researcher was blind to hypotheses. The researcher also recruited passersby verbally, offering them a snack in exchange for their participation. No demographic data were collected to simplify the procedure.

Participants received a laminated instruction card that read, "Your task: Improve this Lego™ project. Please let the researcher know when you are done." If participants requested any additional information about the task, the researcher referred the participant back to the instruction card. The researcher escorted the participant to a semi-private workstation where the Duplo structure to be transformed was sitting on a small desk with the remaining Duplo blocks scattered in a pile off to the side. When the participant returned the transformed structure to the researcher, the researcher counted and recorded the total number of blocks that were added, removed, and moved, and took a photograph of the final structure for record-keeping.

Participants worked on one of two initial structures (versions *A* and *B*), which included 8 and 10 Duplo blocks, respectively, each arranged on an 8 × 8 flat base. Researchers designed both versions with the aim of avoiding any clear pattern or resemblance to familiar objects. Each version of the structure was given to 21 participants.

The observation of interest in study S3 was whether the final structure contained more or fewer blocks than the original structure. One transformation (from across both versions of the original structure) included the same number of blocks as the original and was therefore excluded from analysis. Among the remaining transformed structures, 5 (12%) used fewer blocks than the original, and 36 (88%) used more blocks than the original. A test of whether 5/41 differs from 50% rejects the null hypothesis that people were as likely to create a transformed structure with fewer blocks as to create a transformed structure with more blocks, two-sided binomial distribution probability, $p < .001$.

2.4 Study S4 – Block structures created via randomized process

In the task for study S4, participants transformed a small Lego structure. This allowed observation of additive and subtractive behavior when transforming physical objects, which in this case, were created using a randomized process.

We recruited participants ($N = 60$) in the same manner as study S3. The researcher was again blind to hypotheses. No demographic data were collected to simplify the procedure.

The procedure was the same as study S3 except in study S4 participants worked on one of three initial Lego structures (versions *A*, *B*, and *C*), each of which included 12 Lego blocks, selected from a pool of 24 blocks, arranged on a 6 × 8 flat base. The initial structures were constructed by assigning a reference number to each of the original blocks (1 to 24 in versions A and B, and 1 to 20 in version C), and to each knob on the 6 × 8 flat base (1 to 48). Initial structures were then constructed by iteratively: (i) selecting at random (without replacement) a numbered block from the pile of 24; (ii) selecting at random whether the brick would be oriented horizontally or vertically; (iii) selecting at random (with replacement) a numbered base knob on which to place the top left corner of the selected block. If a block was to be placed on top of an already placed block, it started a new layer. Step (iii) also included the caveat that any brick hanging off the perimeter would be moved "up and left" until it fit. We presented each of the three originals to 20 different participants.

The observation of interest in study S4 was whether the final structure contained more or fewer blocks than the original structure. Across all three versions of the structure, 1 revision (2%)

used fewer blocks than the original and 59 revisions (98%) used more blocks than the original. A test of whether 1/60 differs from 50% rejects the null hypothesis that people were as likely to create a revision with fewer blocks as to create a revision with more blocks, two-sided binomial distribution probability, $p < .001$.

2.5 Study S5 – Block structures created independently

In the task for study S5, participants transformed a small Lego structure. This allowed observation of additive and subtractive behavior when transforming physical objects, which in this case, were created by independent participants, not researchers.

We recruited participants ($N = 20$) in the same manner as in studies S3 and S4. The researcher was again blind to hypotheses. No demographic data were collected.

The procedure was the same as in studies S3 and S4, except in study S5, participants worked on an original structure that had been created by another "builder" participant. These builder participants had been previously recruited using similar means and had received a 20-brick Lego kit (the same composition of bricks as study S4 Version C), and an instruction card that read, "Your task: Create a 10-piece Lego™ project. Please let the researcher know when you are done." Twenty usable original structures were created. (Two originals were discarded because the builder participant interacted with other participants).

The observation of interest in study S5 was whether the final structure contained more or fewer blocks than the original structure. Across the independently created original structures, 1 revision (5%) used fewer blocks than the original and 19 revisions (95%) used more blocks than the original. A test of whether 1/20 differs from 50% rejects the null hypothesis that people were as likely to create a revision with fewer bricks as to create a revision with more bricks, two-sided binomial distribution probability, $p < .001$.

2.6 Study S6 – One's own writing

In the task for study S6, participants edited a piece of their own writing, which allowed an observation of additive and subtractive revisions.

We recruited 268 participants from a participant pool maintained at a large public university in the United States. Among the initial respondents were 167 undergraduate students, 38 graduate students, 24 university community members, and 39 who did not report their affiliation. Of these 268 respondents, we excluded 40 because they did not complete the survey and 30 because they made zero revisions (i.e., ignoring the prompt instructions, or clicking through), yielding a final sample of $N = 198$ participants (68 men, 129 women, 1 unspecified; M_{age} = 21.8, $SD = 6.72$). Participants received either partial course credit or entry into a raffle for a $50 Amazon gift card.

First, participants read a short article adapted from an original news story that described the discovery of King Richard III's bones below a parking lot in Leicester, UK. Instructions asked participants to write a summary of that article, using 6—8 sentences or more. After participants submitted their summary, the next screen in the survey instructed them to "edit your summary with the goal of improving how well you summarized the article." Below the instructions was a copy of the participant's original summary in an editable text box.

The observation of interest in study S6 was whether the transformed summary contained more or fewer words than the original. Seven transformed summaries used the same number of words as the original and were excluded from analysis. Among the remaining summaries, 32 (17%) used fewer words than the original and 159 transformed summaries (83%) used more words. A test of whether 32/191 differs from 50% rejects the null hypothesis that participants were as likely to create a transformed summary with fewer words as to create a transformed summary with more words, two-sided binomial distribution probability, $p < .001$.

2.7 Study S7 – Others' writing

In the task for study S7, participants edited a piece of someone else's writing (a yoked participant from study S6), which allowed an observation of additive and subtractive revisions to other people's ideas.

We recruited 257 participants from the same pool used in study S6. Among the initial respondents were 162 undergraduate students, 51 graduate students, 17 university community members, and 27 who did not report their affiliation. Of these 257 respondents, we excluded 28 because they did not complete the survey, and 21 because they made zero revisions. To maintain independence of observations, we also excluded 8 more respondents because of a procedural error in which we assigned them an original summary that had already been presented to another participant. This yielded a final sample of $N = 200$ participants (58 men; 140 women; 2 unspecified; $M_{\text{age}} = 21.40$, $SD = 6.04$). Participants received either partial course credit or entry into a raffle for a $50 Amazon gift card.

The procedure for study S7 was the same as in study S6, except that after participants submitted their original summary, the next screen in the survey had a copy of *someone else's* original summary in an editable text box.

The observation of interest in study S7 was whether the transformed summary contained more or fewer words than the original. Eleven revised summaries used the same number of words as the original and were excluded from analysis. Among the remaining summaries, 61 (32%) used fewer words than the original and 128 transformed summaries (68%) used more words. A test of whether 61/189 differs from 50% rejects the null hypothesis that participants were as likely to create a transformed summary with fewer words as to create a transformed summary with more words, two-sided binomial distribution probability, $p < .001$.

Across studies S6 and S7, participants were significantly more likely to revise the summaries by adding words than by subtracting them, a finding that seemingly reaffirms the need for Strunk and White's timeless advice to "omit needless words"[3].

2.8 Study S8 – Day-trip itinerary

In the task for study S8, participants transformed a travel itinerary, which allowed an observation of additive and subtractive transformations to one's "situation."

We recruited 211 MTurk respondents. Among these respondents, 7 withdrew before completing the task, 16 did not complete the task, 42 failed an attention check item ("How often do you like to get speeding tickets?")[4], and 2 were from duplicated IP addresses, producing a final sample of 144 participants (85 men, 59 women; $M_{\text{age}} = 37.09$, $SD = 12.81$). Participants received $0.75 for participating.

Participants reviewed a travel itinerary and read the following instructions:

> *Here is an itinerary for a day-long trip to Washington D. C. Your job is to* ***improve the quality of this trip****. Make whatever changes you think are necessary to* ***make it a better trip****. Assume that any extra time will be spent in the hotel room. The itinerary is shown on the right, and the list of possible activities to choose from is on the left. Drag and drop items to make your changes. You may make any changes at all that you think should be made.*

The itinerary displayed in two sections, the first reading "Morning: 8:00am – 3:00pm" and the second reading "Afternoon/Evening: 3:00pm – 10:00pm." Each of these sections on the itinerary was pre-populated with six activities (e.g., "Visit the White House", "See trees at the

National Arboretum", "Go shopping"), for a total of 12 activities on each participant's original 14-hour itinerary. We designed the original itinerary to be exceptionally busy, so that it could potentially be improved by subtracting activities. In addition to the pre-populated itinerary activities, a menu of 16 more potential activities was displayed, including, "Other (Feel free to add your own activity"). Using a drag-and-drop interface, participants could transform their itinerary by rearranging, adding, and subtracting activities.

The observation of interest in study S8 was whether the transformed itinerary contained more or fewer activities than the original itinerary. Forty-three revised itineraries used the same number of activities as the original and were excluded from analysis. Among the remaining itineraries, 28 (28%) included fewer activities than the original and 73 (72%) included more activities than the original. A test of whether 28/101 differs from 50% rejects the null hypothesis that participants were as likely to create a transformed itinerary with fewer activities as to create a transformed itinerary with more activities, two-sided binomial distribution probability, $p < .001$.

2.9 Study S9 – Changing originals with more and less anomalous components

We used a correlational design in study S9 to observe whether participants would be more likely to subtract from stimuli with anomalous components. In the task for study S9, participants transformed grilled cheese sandwich recipes.

Stimuli development. We first developed an external metric of atypicality by examining 120 grilled cheese sandwich recipes (including 994 different ingredients) in user-submitted posts on allrecipes.com. From this recipe data, we created a list of the 22 most common ingredients after bread, cheese, fat (butter or oil), and seasoning or condiments (salt, pepper, other spices, and sauces). To this list of typical ingredients, we added five additional ingredients assumed to be extremely atypical for a grilled-cheese sandwich, such as chocolate.

To validate this list of 27 ingredients, we invited an independent sample of MTurk workers to rate them on atypicality. Among this group of respondents, 10 did not complete the survey, 13 did not give a valid/codable response, and 5 were from duplicated IP addresses, leaving a final sample of 80 participants (51 men, 28 women, 1 unspecified, M_{age} = 34.71, *SD* = 10.92). Participants received payment of $0.75. We asked the raters to imagine going to a diner and seeing on the menu a grilled-cheese sandwich with each of the 27 ingredients (e.g., "grilled cheese with tomato," "grilled cheese with chocolate," and so on). For each of the 27 ingredients, these participants indicated, "how surprised would you be to see a grilled cheese sandwich with [ingredient] on the menu?" (from 1 = *not at all surprised* to 5 = *very surprised*). Thus, we assigned each of the 27 ingredients a mean-atypicality score based on 80 participant ratings.

Main task. For the focal task in study S9, we recruited a separate sample of 335 respondents from MTurk. Among the respondents, 28 did not complete the survey, 8 did not give a valid/codable response, 14 were from duplicated IP addresses, and one failed an attention check, leaving a final sample of 284 (156 men, 126 women, 2 unspecified; M_{age} = 35.62, *SD* = 10.79). Participants received payment of $0.75.

All participants in the main task read, "Imagine you are hungry and decide to make a grilled cheese sandwich for lunch. Below is a recipe for a grilled cheese sandwich." Participants then viewed a list of four ingredients: bread; cheese; butter; and a fourth ingredient, randomly selected from the pre-tested list of 27 (banana: *n* = 7; dill pickles, lime: *n* = 8; anchovies: *n* = 9; apple, chocolate, lemon, lettuce, tomato, vodka, yogurt: *n* = 10; avocado, bacon, bell peppers, egg, liver, mushrooms, oil, onion, turkey, vanilla extract: *n* = 11; broccoli, ham, nutmeg, olives, peanut butter, spinach: *n* = 12). Instructions prompted participants: "In a few sentences, please describe how you would change this recipe when making your grilled cheese sandwich." Responses were coded independently by three raters who counted both the number of ingredients added (α = .94), and the number of ingredients removed (α = .92). Disagreements were handled by averaging the responses of the three coders.

We collected a number of exploratory measures that we do not report on here: "I would add ingredients to the original recipe to improve it;" and "I would remove ingredients from the original recipe to improve it" (*yes* or *no*); "How much would your suggestions change the recipe?" and "How much would your suggested changes improve the recipe?" (1 = *not at all*, 5 = *very much*); and questions probing dietary restrictions and preferences, including "I like to try new foods" and "I like to eat unusual foods" (1 = *strongly disagree*; 5 = *strongly agree*). We also collected demographic information.

Our primary hypothesis was that participants would be more likely to produce a subtractive transformation (i.e., a recipe with fewer than four ingredients) when the fourth ingredient was more atypical. A logistic regression predicting a participant's (N = 284) likelihood of subtracting from the fourth ingredient's atypicality score showed that recipes with atypical ingredients were more likely to elicit subtractive transformations, b = .41, z = 3.05, p = .002 (see Extended Data Figure 1). Analyses treating subtraction as a continuous variable also showed that the atypicality of the fourth ingredient was negatively associated with the total number of ingredients in the transformed recipe, $r(282) = -.25$, $p < .001$.

Additionally, at the level of ingredient (k = 27), an ingredient's average atypicality score positively predicted the percentage of participants who subtracted, $r(25) = .54$, p = .003.

2.10 Study S10 – Changing originals with more and fewer components

We used a correlational design in study S10 to observe whether participants would be more likely to subtract from stimuli with many distinct components. In the task for study S10, participants transformed recipes for vegetable soup.

We recruited 362 respondents from MTurk. Among the respondents, 36 did not complete the survey, 50 did not give a valid/codable response, 21 were from duplicated IP addresses, and 5 failed an attention check, yielding a final sample of 250 (134 men, 108 women, 8 unspecified; M_{age} = 36.28, *SD* = 11.59). Participants received payment of $0.75.

Participants were randomly assigned to transform one of three stimuli: a 5-ingredient vegetable soup recipe including vegetable broth, carrots, peas, garlic, and salt/pepper (*n* = 84); a 10-ingredient vegetable soup recipe additionally including onion, celery, oregano, potatoes, and thyme (*n* = 89); or a 15-ingredient recipe additionally included green beans, corn, zucchini, parsley, and leeks (*n* = 77).

All participants read, "Below you will see a list of ingredients for a soup recipe. Your job is to make any and all changes necessary to improve this soup. Assume that this soup is for someone who has no dietary restrictions or strong food dislikes." Participants were provided with either the 5-, 10-, or 15-ingredient original recipe in an editable text box.

We collected a number of exploratory measures that we do not report on here: "I would add ingredients to the original recipe to improve it;" and "I would remove ingredients from the original recipe to improve it" (1 = *not at all*, 5 = *very much*); "How much would your suggestions change the recipe?" and "How much would your suggested changes improve the recipe?" (1 = *not at all*, 5 = *very much*); and questions probing dietary restrictions and preferences, including "I like to try new foods” and “I like to eat unusual foods” (1 = *strongly disagree*; 5 = *strongly agree*). We also collected demographic information.

In the 5-ingredient condition, 2 participants produced a subtractive transformation, 74 participants produced an additive transformation, and 8 used the same number of ingredients as the original. In the 10-ingredient condition, 31 participants produced a subtractive transformation, 51 participants produced an additive transformation, and 7 used the same number of ingredients as the original. In the 15-ingredient condition, 52 participants produced a subtractive transformation, 16 participants produced an additive transformation, and 9 used the same number of ingredients as the original. As original recipes became more complicated, participants were more likely to produce a subtractive transformation, $\chi^2(4, N = 250) = 83.11$, $p < .001$, Cramer’s $V = .41$ (see Extended Data Table 2).

The total number of ingredients added or subtracted also varied by condition. In the 5-ingredient condition, participants added an average of 2.85 ingredients ($SD = 2.13$), which was significantly greater than zero, one-sample $t(83) = 12.27$, $p < .001$, $d = 1.34$. In the 10-ingredient condition, participants added an average of 0.45 ingredients ($SD = 2.12$) ingredients, which was greater than zero, one-sample $t(88) = 2.00$, $p = .048$, $d = .21$. Finally, participants tended to subtract from the 15-ingredient recipe, ending up with 1.87 *fewer* ingredients on average (SD = 2.53), which was significantly lower than zero, one-sample $t(76) = -6.50$, $p < .001$, $d = -.74$.

Note that, to keep the recipes coherent and the design simple, we allowed the number of ingredients to be fully confounded with the type of ingredients. Therefore, we cannot attribute a condition difference specifically to either variable. For example, the tendency to subtract from the 15-ingredient soup could be a “15-ingredient effect” or a “zucchini effect.” Either way, the result is consistent with the general notion that atypical, disliked, and/or superfluous components are likely to draw more attention to the possibility of subtraction.

2.11 Study S11 – Perceived legitimacy of additive and subtractive transformations

We designed this experiment as a post-hoc assessment of a potential alternative explanation for the results of experiment 1. In experiment 1, participants were more likely to transform a Lego tower by subtraction when instructions included an explicit subtraction cue than when they did not. Noting that instructions in both conditions gave participants blanket permission to alter the structure in any way, we concluded that the subtraction cue increased subtractive transformations because it increased participants' likelihood of thinking of the idea. However, a potential alternative explanation is that the subtraction cue increased subtractive transformations because it increased participants' judgments of subtraction's legitimacy. In other words, it is possible that a non-trivial number of participants in the control condition thought of subtraction but passed it up in favor of an additive approach because they inferred that subtraction was not allowed; whereas participants in the subtraction-cue condition had been assured that subtraction was allowed because it was mentioned explicitly. To investigate this alternative possibility, we designed a parallel task that held knowledge activation constant (by presenting both transformation possibilities) and measured participants' evaluations of each kind of transformation following one of the two instruction sets. If the difference in experiment 1 is attributable to participants' evaluation of the solutions, rather than to the differential activation of subtractive ideas, then we would expect the instructions to affect evaluations in the parallel set-up of study S11.

Because the goal of the experiment was to determine whether the subtraction-cue (vs. control) instructions legitimized the subtractive transformation in participants' eyes, we needed a paradigm that would allow us to present both kinds of transformations to participants without implying the researchers' endorsement of either one. We created a minimal managerial scenario, telling participants that they were the manager and that their team had come up with "two possible ideas" in response to the task specifications from experiment 1. Using this approach, we could deliver the same instructions, show the same Lego setup, offer the same contingent bonus, and solicit participants' evaluations of both the additive and subtractive transformations, without representing either approach as approved or preferred by the researchers. We designed study S11 for online administration and restricted recruitment to a psychology department participant pool (instead of in-person administration with the broader university community as in experiment 1) because it occurred during a time of in-person data collection restrictions due to the Covid-19 pandemic. Both studies were conducted at the same large public United States university. Based on the sample size used in experiment 1, we aimed to recruit 200 participants. Among the respondents, 1 reported technical difficulties; all other participants were included in the analysis, yielding a final sample of $N = 210$ (49 men, 160 women, 1 non-binary/third-gender; $M_{\text{age}} = 18.69$, $SD = 1.23$). Participants received course credit and any earned bonus payment.

The procedure mimicked that of experiment 1 as closely as possible, with the addition of the managerial scenario and necessary adjustments for online administration. After the consent form, the first page of the online survey read:

> *Today, you are the manager of a small team. Your team's task is to change a structure made of Lego bricks so that it can hold a real brick over an action figure's head. Your bonus payment will depend on performance. [This is based on a real task we have done in-person. Like the participants who completed the task in-person, you have the opportunity to earn a cash bonus based on your performance.]*

From there, the survey described the original experiment 1 set-up as closely as possible. Participants read the same task overview and saw photographs of the original Lego structure from four different angles. Instructions said, “On the table there is also a small pile of Legos, a regular-sized masonry brick (approximately 3lbs.), and an instruction card.” We presented an image of the same card we delivered to participants in experiment 1. As in experiment 1, participants were randomly assigned to either a control version of those instructions (“each brick you add costs ten cents”; n = 106) or a subtraction-cue version of those instructions (“each brick you add costs ten cents, but removing bricks is free and costs nothing”; n = 104).

Next, participants read that “your team came up with two possible ideas.” We showed two short gifs, which depicted a closeup of a person’s hands producing each kind of transformation (in counterbalanced order), either adding 3 bricks as support pillars to the precarious platform or removing the 1 extraneous brick to stabilize it. Below the looping gifs, the survey posed three questions:

> *Is Idea #1 a valid solution (In other words, does it accomplish the task without breaking any rules?)* [response options: *yes* or *no*]
>
> *Is Idea #2 a valid solution (In other words, does it accomplish the task without breaking any rules?)* [response options: *yes* or *no*]
>
> *As the manager, you get final say in which solution to pursue and should choose the BEST solution. Which solution would you choose?*” [response options: *Idea #1* or *Idea* #2]

Most participants saw the subtractive transformation as valid regardless of whether the subtraction cue was present in the instructions they read (100%) or not (98%). (Most participants also saw the additive transformation as valid regardless of whether the subtraction cue was presented to them (95%) or not (99%).) Given that experiment 1 and study S11 presented participants with the same instructions, the comprehension check in study S11 can be taken as suggestive evidence that participants in experiment 1 who generated their own subtractive transformations would have been unlikely to reject them as illegitimate based on the instructions.

There was not evidence that the cue affected participants' likelihood of choosing the subtractive transformation as the best one when both options were presented to them: 86% of participants chose the subtractive approach in the cue condition and 82% chose it in the control condition, $\chi^2 = 0.47$, $p = .491$, $\varphi = 0.05$. The 95% CI of the observed effect size [-.09, .17] did not include the observed effect size ($\varphi = 0.20$) of the cue manipulation on subtractive transformations in experiment 1. Study S11 failed to find support for the choice explanation of the cue manipulation but lacks the statistical power to rule out effect sizes smaller than $\varphi = .18$.

2.12 Study S12 – Overlooking subtraction across cultures

Study S12 probed whether participant samples outside the United States systematically overlook subtraction. We recruited 239 participants from universities in the United States, Germany, and Japan. Among the respondents, 13 did not complete the task and one participant provided a duplicate response, yielding a final sample of 225 (students in Japan = 48, students in Germany = 78, students in the United States = 99; 64 men, 159 women, 2 unspecified; M_{age} = 22.61, SD = 7.43). Participants received either: entry into a lottery for a ¥400 amazon.co.jp gift card (Japanese sample); entry in a lottery for one of three €50 amazon.de gift cards (German sample); or course credit (US sample).

Participants first indicated whether they lived in Japan, Germany, or the United States. (Translated study materials were produced by a native German- or Japanese-speaking professional with a background in psychology research). Participants then completed a single-trial version of the grid task used in experiments 6—8.

Participants transformed one of two different grid patterns for which the correct transformation could only be achieved by removing squares from the grid. Participants then viewed an image of a new pattern with two possible transformations labeled “A” and “B”: one of which was an incorrect additive transformation and the other of which was a correct subtractive transformation. The placement of the additive versus subtractive transformation as position “A” versus “B” was counterbalanced across participants. Participants indicated, “Which of these two solutions is the best one, according to the instructions for the task?” Participants selected one of three options, “Solution A is the best”, “Solution B is the best”, and “Solution A and B are equally good.” We coded participants’ responses as correct if they identified the subtractive transformation as the best, and incorrect if they selected either of the other two options.

Our dependent measure of interest was the percentage of participants who initially produced the additive transformation and then subsequently identified the subtractive transformation as correct in a multiple-choice format. Across all participants, including the German, Japanese, and United States university samples, 93% identified the subtractive transformation as correct, whereas 77% provided that transformation initially, McNemar’s $\chi^2(1, N = 225) = 25.69$, $p < .001$.

As summarized in the main text (and Extended Data Table 3), 90% of participants in the combined German and Japanese sample eventually recognized the (correct) subtractive transformation and 74% initially produced it, McNemar’s $\chi^2(1, N = 126) = 13.79$, $p < .001$. Low power prevents confident interpretation of any difference between the German and Japanese samples. Note, however, that in the German sample, 91% recognized the (correct) subtractive transformation and 67% produced it, McNemar’s $\chi^2(1, N = 78) = 14.09$, $p < .001$). In the Japanese sample, 90% recognized the (correct) subtractive transformation and 85% produced it (McNemar’s $\chi^2(1, N = 48) = .17$, $p = .683$). Thus, while combining the subsamples indicates that subtraction oversight is widespread, additional data would be needed to answer finer-grained questions about any cultural differences.

Supplementary Information References

1. W. Viechtbauer, Conducting meta-analysis in R with the metafor package. J. of Stat. Software., 36, 1-48 (2010).
2. Johnson, E. J., Häubl, G., & Keinan, A. Aspects of endowment: A query theory of value construction. *J. Exper. Psychol.: Learning, Mem., and Cog.* **33**, 461-474.
3. Strunk, W. & White, E. B. *The Elements of Style: Fourth Edition.* (Allyn and Bacon, Boston, MA, 1999).
4. M. R. Maniaci, R. D. Rogge, Caring about carelessness: Participant inattention and its effects on research. J. Research in Person., 48, 61-83 (2014).