**Pay more and save less: Joint evaluation mode causes biased judgments of efficiency upgrades**

People are often faced with opportunities to upgrade their products and services to faster or more efficient ones. People then need to judge the offered upgrade either by itself or in comparison to other upgrades. The evaluability theory argues that separate evaluations of options are often influenced by attributes which are easier to evaluate, even if these attributes are not important. As a result, when given an inferior and superior options to evaluate separately, people show biased judgments and often choose the inferior option. In this paper we show an opposite pattern in the context of efficiency or speed upgrades: When asked to judge between two upgrade offers (for home Internet speed or cars’ fuel-efficiency) people preferred a larger – but less economical –upgrade in a joint evaluation mode, whereas in a separate mode their evaluations were less biased and more calibrated to the actual time or fuel savings. In two experiments, we found that people were willing to pay or to accept offers for larger upgrades more than they should have, compared to their more calibrated preferences when the upgrades were evaluated separately. These findings hold both theoretical implications for the evaluability theory, as well as practical and policy implications about how efficiency upgrade options should be presented to decision-makers.

*Keywords: efficiency savings, time saving bias, the mpg illusion, joint and separate evaluations, evaluability theory*

**Introduction**

People sometimes evaluate one out of two objects better when they are presented jointly compared to when each object is presented separately (Hsee, 1998). For example, people expressed higher willingness-to-pay (WTP) for a 7 oz ice-cream served in a 5 oz cup than they were willing to pay for an 8 oz ice-cream served in a 10 oz cup, when these were presented separately. Hsee (1998) explained this "less is better" phenomenon using *the evaluability theory*: Separate evaluations of options are often more influenced by attributes which are easier to evaluate, and less by attributes that are harder to evaluate, even when the latter might be more important. In the above example, when presented with the ice creams separately, people evaluated the amount of ice-cream relative to the size of the cup, instead of evaluating the ice-cream's amount, which should be regarded as a more important attribute than the cup in which it is served. Indeed, when both ice creams were evaluated jointly, people considered the more important attribute (the amount of ice-cream) and disregarded the size of the cup (Hsee, 1998, 2000).

Biased judgments were also found in separate evaluation (SE) vs. joint evaluation (JE) in other contexts: judging job candidates in SE produced a bias in favor of men (vs. women) that were selected more for math tasks in SE vs. JE (Bohnet, van Green, & Bazerman, 2015). SE was found to accentuate negative features of a product, which were then weighed more than positive features in SE compared to JE modes (Willemsen & Keren, 2004). Kogut & Ritov (2005) found that charitable giving for a single victim exceed contributions for a group when these were judged separately, but preference reversed when one has to choose between contributing to the single individual and contributing to the group. In other words, the identifiable victim effect (Small & Loewenstein, 2003) occurs more in SE than in JE.

Hsee (1996) also showed preference reversals between two objects in SE vs. JE in contexts where neither preference is objectively better than the other. Such pattern was evident in contexts of choices between two objects in which there is an attribute that is more evaluable in SE and another attribute that is more evaluable in JE, however, none of these attributes is more important than the other. Several studies showed the effect of SE vs. JE in other contexts, in which there is no "correct" answer. For example, Davidai and Shafir (2018) showed that people endorse behavioral nudges in SE, however, they prefer more traditional public policies over nudges in JE.

Thus, it appears that, from both a descriptive and prescriptive points of view, JE modes are often better at helping people form more accurate judgments that lead to more consistent choices. Bazerman, Gino, and Tsay (2011) also describe JE to be an effective tool in helping decision-makers achieve unbiased judgments in the contexts of morality judgments. The advantage of JE mode apparently comes from enabling decision-makers to focus on the more important and relevant attributes, and less to irrelevant attributes that become more salient when the choice is judged separately. However, JE might also require more cognitive effort in order to accurately evaluate the two options on the important attribute. In all the cases reviewed above, it does not seem that JE (of ice-cream cups, job candidates or products) should have been more cognitively taxing, and thus it is no surprise that decision-makers performed better in those cases. However, in other cases, where judgments require a higher degree of cognitive effort in order to identify the best choice, it is possible that SE, and not JE, might actually lead people to less biased judgments and decisions. In this paper, we examine the case of judging efficiency upgrades, in which we find that evaluations are indeed biased in JE, but not in SE. Specifically, when people consider upgrading a product or a service (e.g., a car with higher fuel efficiency or an Internet service with higher speed), they show an overestimation of the savings of the larger upgrade relative to a small upgrade, but only when both upgrades are presented in JE. However, less or no bias is found in SE. This bias manifests in people willing to over-pay for the larger (vs. the smaller) upgrade when they are presented jointly, compared to separately. But, because the larger upgrade actually offers a smaller relative saving, this leads people to actually pay more and save less.

In contrast to the previous literature on joint vs. separate evaluations, we offer a different theoretical account for this phenomenon. Previous demonstrations of preference reversals in joint vs. separate evaluations have been explained by the evaluability of different attributes (Hsee, 1996). We suggest that the effect of SE vs. JE in the context of efficiency savings or upgrades could be better explained by the different availability of cognitive strategies (or heuristics) used for judgment of the upgrades' efficiency savings. Specifically, the availability of less accurate heuristics increases when upgrade options are presented jointly, whereas the availability of normative strategies and less biased heuristics increases when upgrade options are presented separately. Thus, the context of efficiency measures presents an interesting special case within the research paradigm on joint vs. separate evaluations. To better define this theoretical account, we first explain the key features of decisions involving efficiency upgrades that distinguish them from other contexts that have been previously studied in the joint vs. separate evaluations paradigm.

**The biased judgments of efficiency upgrades**

Efficiency (or productivity) measures generally present the number of units that can be produced or consumed in a given period of time or amount of a resource. Examples include driving speed (e.g., mph or km/h), manufacturing speed (e.g., products per hour), Internet download speed (e.g, megabits per second), fuel economy (e.g., miles per gallon or liters per km), etc. Although ubiquitous in many domains of our daily lives, research has shown that efficiency measures entail an inherent drawback, as they seem to hamper people’s ability to estimate savings and cause them to make sub-optimal decisions (Larrick, Soll, & Keeney, 2015). Presented with information about the efficiency of products or services, people have been found to overestimate savings when performed from relatively high initial values. For example, people falsely believe that upgrading a car from 20 to 40 MPG would save more fuel (for a given distance) than upgrading a car from 10 to 15 MPG, even though the latter actually saves more (25 vs. 33 liters per 1,000 miles; Larrick & Soll, 2008). Similarly, people falsely believe that increasing speed from 40 to 60 mph would result in saving more time than increasing speed from 20 to 30 mph when the opposite is correct (1 vs. 2 hours per 120 miles; e.g., Peer & Gamliel, 2013; Svenson, 2008). Similar biases were found in studies using a driving simulator (Eriksson, Svenson, & Eriksson, 2013), and in other contexts such as savings of patients’ wait time (Svenson, 2008), estimating a productivity increase in a manufacturing line (Svenson, 2011), and judging products’ speed (e.g., Internet services or printers) and their related quality (De Langhe & Puntoni, 2016).

These biases are caused by replacing the normative evaluations, that require calculating the savings, with cognitive shortcuts (i.e., heuristics). Previous research demonstrated that people apply several heuristics when evaluating savings, and that they apply simpler heuristics that cause larger biases as the task becomes more complicated. In the context of time or fuel savings, the evaluation mode (joint vs. separate) can influence people’s decision by leading them to choose different strategies, or heuristics, for the judgment of the time or fuel saved when upgrading speed or fuel efficiency. Previous studies identified two main non-normative cognitive strategies (heuristics), that characterize people’s biased judgments. One heuristic includes favoring the speed/efficiency increase that has the higher absolute *Difference* between the initial and increased values (Peer & Gamliel, 2012). Another heuristic uses the *Proportional* increase, computed from the increased value (Svenson, 2008). For example, when judging whether more time is saved when A) increasing from 30 to 40 kph vs. B) from 60 to 80 kph, followers of the *Difference* heuristic would opt for option B, whereas followers of the *Proportion* heuristic would judge both options equally. Although both the Difference and the Proportion heuristics lead to biased estimations (option A actually saves more time than option B in the above example), the use of the *Difference* heuristic typically results in more biased judgments relative to the use of the *Proportion* heuristic (Gamliel & Peer, 2019). Previous studies have also suggested that the type of task presented to participants affects the heuristic they use. Specifically, when asked to estimate saving of an efficiency upgrade, more people used the *Proportion* heuristic, but when asked to estimate savings of several upgrades simultaneously the use of the *Proportion* heuristic decreased in favor of the *Difference* heuristic (De Langhe & Puntoni, 2016; Gamliel & Peer, 2019).

**Efficiency upgrade judgments in joint vs. separate evaluation modes**

Although it has not been tested before, one can assume that evaluating two upgrade options in a JE mode would be more cognitively taxing than evaluating only one in a SE mode. This should especially be the case for judgments that require some mathematical calculations, as in the examples of estimating time savings when increasing speed, or fuel savings when upgrading cars. Thus, we predicted that faced with upgrade options in a SE mode, people would estimate the savings using either the normative formula or use more complex heuristic (i.e., the *Proportion* heuristic), resulting in small to no bias in judging the larger and smaller upgrades and expressing calibrated WTP. But, when faced with upgrade options simultaneously in a JE mode, people might abandon both the normative formula and the extra step of converting upgrades to proportions, and follow the simpler *Difference* heuristic. As the use of the Difference heuristic typically results in more biased evaluations, people are thus expected to evaluate the saving of the larger upgrade in JE mode in a more biased manner relative to SE.

For example, assume an office manager who is considering the price he is willing to pay for upgrading the office printer that currently prints 50 pages per minute (PPM). He has an option to upgrade to Printer X that prints 100 PPM. If asked how much he is willing to pay for Printer X, he might estimate that printing 500 pages a day using the current printer would require ten minutes relative to five minutes using Printer X, and estimate that this daily 5-minutes time saving justify paying $W. Now imagine there is an alternative option to upgrade to Printer Y, which prints 150 PPM (assuming all other remains equal). Printing 500 pages using Printer Y would require 3.33 minutes (500 pages / 150 PPM = 3.33 minutes), and this daily 6.67 minutes time-saving might justify paying $V. As the latter upgrade saves 6.67 minutes and the former only 5 minutes, the estimation of $V should be 1.33 times higher than $W. We would thus expect, for example, that if the office manager was willing to pay $100 for upgrading to Printer X, he would be willing to pay approximately $133 for upgrading to printer Y. That is, the ratio of $V/W$ should correspond to the ratio of the printing speed of Printer Y and the printing speed of Printer X.

However, in a joint evaluation task, the office manager faces more data and more calculations for establishing his WTP for Printers X and Y. The task of estimating time savings in both upgrades would be more complicated, and the office manager might be inclined to use a simpler heuristic. The *Difference* heuristic might result in the following line of reasoning: If one is willing to pay $W for a 50 PPM difference in the first upgrade (100 PPM vs. 50 PPM), then one should probably be willing to pay **twice** as much for a 100 PPM difference in the second upgrade (150 PPM vs. 50 PPM). Alternatively, using the *Proportion* heuristic might result in the following line of reasoning: If one is willing to pay $W for the first upgrade that doubles the speed (100 PPM vs. 50 PPM), one should probably be willing to pay **1.5 times** to triple the speed (150 PPM vs. 50 PPM). As the larger upgrade actually saves 1.33 of the time saved by the smaller upgrade, both heuristics cause biased estimations, and the *Difference* heuristic results in a larger bias.

**The current research**

We thus predicted that given a smaller upgrade and a larger upgrade in SE mode, people would calculate the savings normatively or use the *Proportion* heuristic, resulting in small to no bias in judging the larger upgrade and expressing WTP. Importantly, when presented with both upgrades in JE mode, people might use a simpler heuristic (e.g., the *Difference* heuristic), and evaluate the relative savings of the two upgrades, resulting in biased evaluations of the larger upgrade's saving and expressing willingness to over-pay for it.

We examined our hypothesis in two experiments. Study 1 presented participants with four vignettes of smaller and larger upgrades relating to speed or fuel efficiency in either SE or JE modes, and participants indicated their WTP for each upgrade. We predicted that the relative difference in WTP between the smaller vs. larger upgrades would be increased in JE vs. SE. This prediction was based on the assertion that when judged jointly the larger upgrade would appear more attractive in terms of its time/fuel savings. Study 2, which was pre-registered, used a forced-choice paradigm, instead of WTP elicitation, to show that people indeed consider the larger upgrade more attractive when presented in JE mode.

**Study 1**

**Method**

**Participants.** We recruited 298 participants from Prolific Academic (*M*age= 34.4; *SD*age = 12.9; 50% were female). Participants were pre-screened to include only U.S. residents, aged 18 or above, who speak English as a first language. In order to maximize participants' attention and to screen out participants who did not read the instructions carefully, we included an attention-check question asking participants what is the topic of the study. In the preceding text participants were instructed to choose the "other" option and state the word "attention" in a box next to it. 38 (13%) of participants failed this question and were omitted from further analyses.

**Design and procedure.** The study was presented as a study on consumers' preferences. Participants read four scenarios, containing offered upgrades of two services (Internet service provider and cars' fast lane) and two products (a printer and a car). In each scenario, participants were presented with initial values and were asked to indicate how much they would be willing to pay for an upgrade to a higher value. Participants were randomly allocated to one of three experimental groups: two groups had a SE task and one had a JE task. For example, the Internet speed scenario in the JE condition read:

Alex and Taylor are both customers of two large Internet Service Providers in the U.S. Both pay a similar monthly fee for home Internet service at a speed of 25 Mbps. Alex's provider offers an option to upgrade Alex's Internet service from 25 Mbps to 50 Mbps. Taylor's provider offers an option to upgrade Taylor's Internet service from 25 Mbps to 100Mbps. How much, in your opinion, should each of them be willing to pay more for their monthly service after the upgrade? What should be the increase - in USD - in Alex's and Taylor's monthly payment compared to the current plan's rate?

The other scenarios asked about WTP for the usage of a fast lane with higher driving speeds (from 40 to 60 or 80 mph), for the purchase of a printer that prints more pages per minute (50 to 100 or 150 pages-per-minute), or for the lease of a car that has a higher fuel efficiency (10 to 20 or 30 MPG). In the separate evaluation conditions, participants were presented with either the small or the large upgrade for each scenario. Participants indicated their WTP for each scenario using a slider that ranged from 0 to $50, $100, $200, or $5000, for the fast lane, Internet service, printer, and car, respectively. The order of the four scenarios was randomized within each experimental group. At the end of the study participants reported their age, gender and any comments they may have had.

## Results and Discussion

Among participants in the JE condition, we omitted from the analyses six responses that included a higher estimate for the smaller upgrade relative to the larger upgrade in a certain scenario. Additionally, we omitted 16 of responses that were over three standard deviations above the mean.

To examine our hypothesis that people would be willing to pay more for the larger upgrade, compared to the smaller one, in joint vs. separate evaluations mode, we computed the difference (in percentages) of WTP between the large and small upgrades for each evaluation condition in each scenario. We compare these ratios of WTP to the normative ratio that is expected from the savings that can be gained (in time or fuel) between the upgrades. Figure 1 shows that the difference in WTP ratio was consistently high when the evaluations of both upgrades was made jointly, and considerably smaller when the upgrades were evaluated separately. For example, when asked for their WTP to upgrade their home Internet speed from 25 to either 50 or 100 Mbps, participants in the JE mode were willing to pay twice as much (100% more) for the larger upgrade (M=$32.26 vs. $16.26, SD=19.7, 11.1). In contrast, participants in the SE mode were willing to only add 27% to the cost of the larger upgrade (M=$21.97 vs. $17.33, SD=12.2, 9.5). This means that the JE mode increased the relative WTP for the large upgrade by 3.7 times. As can be seen in Figure 1, an even larger increase – of 7 times more – was seen for upgrading printers, whereas smaller, but still considerable increases – of 2.17 and 1.8 times – were found for upgrading cars and increasing driving speed, respectively. Table 2 provides all mean WTP for all conditions.

Figure 1. Percentage difference (Ratio) in WTP between large vs. small upgrades in separate and joint evaluations for each scenario in Study 1 (error bars indicate 90% confidence intervals).

Table 2. Mean WTP (and SDs) for the lower and larger upgrades in joint vs. separate evaluations for the scenarios.

|  |  |  |
| --- | --- | --- |
| Scenario | Separate evaluation | Joint evaluation |
| Small upgrade | Large upgrade | Small upgrade | Large upgrade |
| Internet  | 17.33 (9.51) | 21.97 (12.22) | 16.10 (11.10) | 32.26 (19.69) |
| Driving  | 5.69 (3.88) | 8.57 (6.92) | 6.18 (4.78) | 11.88 (8.32) |
| Printer | 33.92 (24.63) | 37.25 (27.85) | 35.41 (26.20) | 60.34 (39.36) |
| Car | 1531 (1016) | 2068 (1454) | 1211 (783) | 2127 (1165) |

Figure 1 further shows that these differences stem from participants in the JE modes providing exaggerated prices compared to the normative standard, which is based on the actual savings that could be gained by the upgrades. For example, in the Internet speed scenario, JE participants’ WTP ratio was twice as large as it should have been, had they followed actual savings. In all scenarios, we found that JE participants’ WTP ratio was significantly larger than the normative ratio, suggesting they overestimated the savings that could be gained by the large upgrade in all cases. Table 2 indeed shows that the difference in WTP between conditions was mainly in the WTP for the large, and not the small upgrades, further corroborating this conclusion. In contrast, the ratio given by participants in the SE condition was much more calibrated and in close resemblance to the normative ratio. It was only significantly different in the printing speed case, and even in that case, SE participants produced a lower, and not higher, ratio of WTP between the upgrades, meaning they did not overestimate the savings of the large upgrade in any of the cases.

To summarize, we found that JE mode caused participants to overestimate what they can gain from a larger upgrade, which led them to provide upward biased WTP for that upgrade and, in essence, be willing to pay much more in order to save much less than they could expect to save. To illustrate this point even further, in order to actually achieve the savings in download time when upgrading Internet speed that could be justified by the WTP ratio given in the JE mode, the large upgrade should have been from 25 Mbps to 200 Mbps (and not 100). This suggests that, de facto, the JE mode made people pay twice as much for getting half as much, for increasing Internet speed, and with similar patterns in the other cases.

However, the fact that people would express higher WTP for a large upgrade in JE mode does not necessarily mean that they would be willing to accepts higher priced offers when asked directly. First, research has long shown that people can express different preferences when asked for WTP vs. choice (e.g., Lichtenstein & Slovic, 1971). Second, people might not hold clear preferences about some of the products or services we examined in Study 1, and this is indeed evident in the relatively large variance observed for WTP. This high variance also made some of our findings not totally conclusive, as the differences in the ratio of WTP in the scenarios of driving speed and fuel efficiency were not statistically significant (evident by the overlap of their confidence intervals). We thus decided to try and replicate our findings using a choice-based paradigm that presents participants with pre-priced upgrade options to which they can accept or reject to express their preferences. We again hypothesized that more people would be willing to accept the more expensive offer for the larger upgrade when it is given in JE vs. SE modes.

**Study 2**

**METHOD**

**Participants.** We recruited 614 participants through the online platform Prolific Academic (*M*age= 37.7; *SD*age = 13.1; 49% were female). Participants were pre-screened to include only U.S. residents, aged 18 or above. As in the previous experiments, we screened out the participants who failed the attention-check question (52 participants; 8.5%).

**Design, materials, and procedure.** Participants were randomly allocated to one of the three experimental conditions: joint condition and two separate conditions of either small or large upgrades. All participants were presented with the Internet and fuel-efficiency scenarios adapted from Study 1. However, instead of asking for WTP for the small or large upgrades, we presented participants with a price for each upgrade and asked them to indicate whether they would be willing to accept that offer or not. In the Internet service scenario participants were given offers to upgrade their home Internet speed from 25 Mbps to 50 or 100 Mbps for $17 or $30, respectively. In the car's fuel efficiency scenario, participants were asked to imagine they had a car on a leasing deal from their workplace, which includes all expenses except for fuel costs, and are offered to upgrade to a different car with higher fuel efficiency of 20 or 30 MPG (from the current 10 MPG) for adding $80 and $120, respectively, to the monthly leasing cost. In the SE conditions, participants were asked to make a choice whether to accept the upgrade or not, whereas in the JE condition they were asked to choose between the two upgrades (assuming they want to upgrade the speed/fuel-efficiency). We included an additional attention-check question asking participants about the values they were presented within the Internet scenario, in order to ensure they were attending the instructions throughout the experiment. Out of the 562 participants who passed the preliminary attention-check question, 462 (85%) passed the second attention-check question, and the following analyses relates to their responses. Other instructions and procedure resembled the ones detailed in Study 1. The study’s hypotheses, methods and analyses were preregistered at <https://aspredicted.org/aw3s2.pdf>.

## Results and Discussion

For the Internet speed upgrade, 85 out of 177 (48%) participants in the lower upgrade condition chose to upgrade the speed, whereas 97 (out of 188; 52%) chose to do so in the higher upgrade condition. In contrast, and as predicted, in the JE condition most participants chose the higher upgrade (97 out of 143; 68%). A similar pattern was found for the fuel-efficiency scenario: 50% (89 out of 177) and 49% (92 out of 188) chose to upgrade the lower and higher fuel-efficiency (respectively), whereas a majority (80 out of 143; 56%) chose the higher upgrade in the JE condition. As can be seen in Figure 2, although the percentages of participants choosing to upgrade the efficiency in the two SE conditions were similar for both the lower and higher upgrade (around 50%), most participants in the JE condition (56%-68%) chose the higher upgrade, as was predicted.

Figure 2. Percent of participants accepting the offer for a small/large upgrade in joint vs. separate evaluation mode in the two scenarios of Study 2.

|  |  |
| --- | --- |
|  |  |

A Z-test for proportions showed that the difference in accepting the large upgrade between the JE and SE conditions in the Internet scenario was statistically significant, Z = 2.97, *p* < .01. However, the difference in the MPG scenario was not statistically significant, Z = 1.26, *p* = 0.1. But, these tests do not account for the base rate of how many people were willing to pay for the small upgrade. Thus, we calculated a difference-in-difference score which showed the difference in WTP for the large vs. small upgrade in the JE condition vs. the difference between the two upgrades in the SE conditions. For the Internet scenario, that difference was (68%-32%)-(52%-48%)=32%, and for the MPG scenario the difference was (56%-44%)-(49%-50%)=11%. A Z-test for these proportions showed significance for both the Internet and the MPG scenarios, Z = 9.64, 5.62, *p* < .01.

These findings corroborate those of Study 1 and show that also when asked to make a choice, and not only express WTP, people tend to prefer larger upgrades in speed or efficiency when these are presented jointly, compared to a separate mode of presentation. Apparently, when upgrades are presented jointly, people are more attracted to judge them using a relative comparison. For example, when asked to choose between upgrading from 10 to 20 or 30 MPG, people, intentionally or not, focus on the proportional difference between the upgrades, judging that the large upgrade offers twice the increase of speed compared to the small one. Following this judgment, people then judge the prices of the upgrades accordingly, and are willing to pay about twice as much for the large vs. small upgrades. Indeed, we designed our study specifically that the price for the large upgrade was a little smaller than twice the price of the small upgrade ($30 vs. $17 in the Internet speed scenario) so that it would be more attractive when judged jointly. But, when these prices were given separately, this relative judgment process was impossible and people did not express a higher preference for the larger upgrade, further corroborating our theory that JE produced the observed bias.

**General Discussion**

The findings of both studies showed that JE cause people to be willing to overpay for larger efficiency upgrades of product or service, whereas no bias was found in SE mode. Our findings differ from those of previous studies on SE vs. JE which typically showed people to be less biased in JE situations (e.g., Bazerman et al., 2011; Bohnet et al., 2015; Hsee, 1996, 1998, 2000; Kogut & Ritov, 2005).

Hsee (1996) proposed the evaluability theory to account for previous demonstrations of preference reversals in joint vs. separate evaluations. The evaluability theory suggests that in separate and joint evaluations of options people might use different attributes because of their different evaluability: some attributes are easier to evaluate in SE mode whereas others are easier to evaluate in JE mode (Hsee, 1996). The theoretical account we propose for the tendency to over-pay for larger efficiency upgrades in JE but not in SE mode differs from the evaluability theory. In the context of efficiency measures (e.g., fuel efficiency and speed as a measure of time efficiency), we propose that the evaluation mode affects the availability of cognitive strategies (e.g., heuristics). Joint evaluation is more complex than separate evaluation, and the differential complexity affects people's choice of the cognitive strategy they use to evaluate the savings, and provide their resulting WTP. Previous studies demonstrated that faced with efficiency saving tasks, people use more simplified heuristics, that result in more biased estimations, as the task becomes more complex (De Langhe & Puntoni, 2016; Gamliel & Peer, 2019). Consistent with these findings, we found that people express more willingness to overpay for larger upgrades of products and services in a more complex task of joint evaluation, but less so in the simpler task of separate evaluation. These findings are consistent with the theoretical account of different availability of cognitive strategies (e.g., heuristics) in the two modes: faced with upgrade options in a SE mode, people estimate the savings following either the normative formula or use more complex heuristic (i.e., the *Proportion* heuristic), resulting in small to no bias in judging the larger and smaller upgrades and expressing calibrated WTP. However, when faced with upgrade options simultaneously in JE mode, people abandon both the normative formula and the extra step of converting upgrades to proportions, and seem to follow the simpler *Difference* heuristic.

Tversky and Kahneman (1974) have already showed that faced with difficult cognitive decisions, people prefer to use cognitive shortcuts (i.e., heuristics) rather than normative rules. In this paper we further suggest that cognitive difficulty can also affect people's choice of heuristics: when the task at hand gets more cognitively difficult and demanding, people revert to simpler heuristics over more complicated ones. We presented the effect of cognitive difficulty on heuristic choice in the special context of efficiency measures, but future research could examine similar effect of reverting to simpler heuristics in other judgment and decision making contexts that involve a choice between heuristics that differ in their complexity.

The current research also contributes to the literature on biases in judgment and decision making in the context of efficiency (or productivity) measures. Our findings are consistent with previous ones suggesting that people might use heuristics rather than normative strategies when trying to estimate time savings (Svenson, 2008), or fuel savings (Larrick & Soll, 2008). Previous research demonstrated biases in the context of efficiency measures in two main contexts – when people are required to estimate the average of efficiency measures (e.g., Gamliel & Peer, 2017; Falk & Lann, 2008), or when they are required to evaluate efficiency savings (Larrick et al., 2015; Svenson, 2008). With respect to efficiency savings, previous studies showed that people overestimate savings when performed from high initial values of fuel efficiency (Larrick & Soll, 2008), and time efficiency (i.e., speed; De Langhe & Puntoni, 2016; Eriksson et al., 2013; Peer & Gamliel, 2013; Svenson, 2008, 2011). Our findings suggest that people's judgments and decisions are also biased when smaller and larger upgrades from a single initial value are evaluated. We propose that in such contexts the willingness to over-pay for the larger efficiency upgrade in JE mode but not in a SE mode could be attributed to the differential availability of the normative rule, the *Difference* heuristic, and the *Proportion* heuristic.

Our findings also offer practical importance as we demonstrated that joint evaluations of smaller and larger upgrades cause people to be willing to overpay for larger upgrades. Consumers’ awareness of their biased WTP in these contexts might assist them in lessening their susceptibility for this bias. Overpaying for larger efficiency upgrades of products and services not only harms the consumer in the short term, but can also lessen firm innovation and adversely affect market efficiency in the long term (De Langhe & Puntoni, 2016). Thus, policy-makers and regulators might benefit from understanding how evaluation mode affects different types of decisions and design policies that could ensure consumers are given the more appropriate and less biased evaluation mode.

**References**

Bazerman, M. H., Gino, F., Shu, L. L., & Tsay, C. J. (2011). Joint evaluation as a real-world tool for managing emotional assessments of morality. *Emotion Review, 3*(3), 290-292.‏

Bohnet, I., Van Geen, A., & Bazerman, M. (2016). When performance trumps gender bias: Joint vs. separate evaluation. *Management Science*, *62*(5), 1225-1234.

Davidai, S., & Shafir, E. (2018). Are ‘nudges’ getting a fair shot? Joint versus separate evaluation. *Behavioural Public Policy*.

De Langhe, B., & Puntoni, S. (2016). Productivity metrics and consumers’ misunderstanding of time savings. *Journal of Marketing Research*, *53*(3), 396-406.

Eriksson, G., Svenson, O., & Eriksson, L. (2013). The time-saving bias: Judgements, cognition and perception. *Judgment and Decision Making, 8*(4), 492-497.‏

Falk, R., & Lann, A. (2008). The allure of equality: Uniformity in probabilistic and statistical judgment. *Cognitive Psychology, 57*(4), 293-334.

Gamliel, E., & Peer, E. (2017). The average fuel‐efficiency fallacy: Overestimation of average fuel efficiency and how it can lead to biased decisions. *Journal of Behavioral Decision Making, 30*(2), 435-445.‏

Gamliel, E., & Peer., E. (2019). Effects of task form on chosen heuristics: Evidence from energy and time-saving decisions. Paper presented at the SPUDM conference, Amsterdam, August 2019.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*, 247-257.

Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making, 11*, 107–121.

Hsee, C. K. (2000). Attribute evaluability and its implications for joint-separate evaluation reversals and beyond. In D. Kahneman & A. Tversky (Eds.), *Choices, Values, and Frames* (pp. 543-563). New York: Cambridge University Press.

Kogut, T., & Ritov, I. (2005). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, *97*(2), 106-116.

Larrick, R. P., & Soll, J. B. (2008). The MPG illusion. *Science, 320*(5883), 1593-4.

Larrick, R. P., Soll, J. B., & Keeney, R.L. (2015). Designing better energy metrics for consumers. *Behavioral Science & Policy, 1*(1), 63-75.

Peer, E., & Gamliel, E. (2012). Estimating time savings: The use of the proportion and percentage heuristics and the role of need for cognition. *Acta Psychologica, 141*(3), 352-359.‏

Peer, E., & Gamliel, E. (2013). Pace yourself: Improving time-saving judgments when increasing activity speed. *Judgment and Decision Making, 8*(2), 106-115.‏

Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty, 26*(1), 5-16. Svenson, O. (2008). Decisions among time saving options: When intuition is strong and wrong. *Acta Psychologica*, 127, 501-509.

Svenson, O. (2011). Biased decisions concerning productivity increase options. *Journal of Economic Psychology, 32*(3), 440-445.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science, 185*(4157), 1124-1131.‏

Willemsen, M. C., & Keren, G. (2004). The role of negative features in joint and separate evaluation. *Journal of Behavioral Decision Making, 17*(4), 313-329.