# Exploring the Decision component of the Activation-Decision-Construction-Action Theory for gain and loss facing scenarios

**Tei Laine (teihlaine@gmail.com)**

**Tomi Silander (tomi.silander@naverlabs.com)**
NAVER LABS, 6 Chemin de Maupertuis
38240 Meylan, France

**Abstract**

Inspired by Masip et al.'s (2016) test of ADCAT model's decision component, we wanted to see if we could replicate their findings using different data from a similar scenario-based study. They found that expected value of telling the truth predicted the decisions to lie or tell the truth more accurately than the expected value of lying, and even better than the motivation to lie, which they defined as a difference between these two expected values. In contrast, in our modeling study the motivation to lie was the best predictor of choices for both actual liars and truth tellers in conditions involving gains and large losses, whereas only in the condition involving large losses the expected value of telling the truth outperformed the expected value of lying. We conclude that whether the participants could gain something or avoid losing something by deceiving determined if they focused on benefits of lying or costs of telling the truth.

**Keywords:** Deception; Motivation; Social cognition; Social-cognitive theory; Risk taking; Lie aversion

## Introduction

Not all deception is bad or reprehensible, or socially unacceptable; people lie in various ways and for various reasons which can be pro-social (e.g., to avoid conflict or help someone), or selfish but do not hurt others (e.g., saving face or making an impression) (dePaulo, Kashy, Kirkendol, & Wyer, 1996; Erat & Gneezy, 2012; Gupta, Sakamoto, & Ortony, 2013). At the same time, the costs and benefits that drive decisions to lie in everyday life may reside in different domains, for instance they can be financial, material, reputational, or psychological (Sakamoto, Laine, & Farber, 2013). Alternatively, they may lie in the same domain and even carry the same absolute (objective) value, but differ in psychological (subjective) value. For example, getting an undeserved discount on a purchase and avoiding paying for damages one has caused may both carry an equal (monetary) value, but most people may choose to deceive in one situation but not in the other; they may be able to justify that deceiving for the discount is only fair ("The product is overpriced, anyway"), but feel good about admitting the guilt of accidentally breaking something and paying for the damages.

Standard economic theories of rational behavior approach decisions to lie as cost-benefit analysis. They posit that whenever the expected benefit from lying exceeds the outcome of being honest a selfish individual, a "homo economicus," should lie, and the decision should be determined solely by the trade-off between the gain from lying and the penalty incurred if detected (Abeler, Becker, & Falk, 2014). Along these lines, Levine, Kim, and Hamel (2010) posit that people "lie for a reason"; they tell the truth if it does not prevent them from attaining their goals, and only when it does they may consider deception. In other words, people lie only when it is more beneficial or less harmful with respect to their goal attainment than truth. Relatively recently economists have started to acknowledge that the act of lying may have an intrinsic cost that deters people from lying even if it would be beneficial (Gneezy, Rockenbach, & Serra-Garcia, 2013).

However, even for social goals people lie a lot less than the economic models predict, and the discrepancy cannot be solely explained by unusually strong risk aversion or pure lie aversion (Dhami & al-Nowaihi, 2007; Gneezy et al., 2013; López-Pérez & Spiegelman, 2012). Cappelen, Sørensen, and Tungodden (2013) showed that non-economic aspects of a choice situation have role in decisions to lie. For instance, in the context of taxes, emotions have been found to influence the propensity to take risk and evade, so that the act of tax reporting may elicit anticipated emotions of how one would feel if audited and punished, and consequently these emotions drive behavior and future tax compliance (Coricelli, Rusconi, & Villeval, 2014). Furthermore, Maciejovsky, Schwarzenberger, and Kirchler (2012) have argued that emotions can moderate the relative effectiveness of economic variables such as audit probabilities and fines in tax ethics.

Walczyk, Harris, Duck, and Mulay (2014) have proposed a quasi-rational model of deceptive decision making, called *Activation-Decision-Construction-Action Theory (ADCAT)*, that combines costs and benefits from different domains, such as material outcomes, and affective responses ranging from apprehension of being detected to thrill of successfully deceiving. According to the model, after considering utilities and probabilities of the most important consequences of all choice options, a decision maker chooses the option that best achieves her goal.

Using everyday scenarios to test ADCAT model's decision component, Masip, Blandón-Gitlin, de la Riva, and Herrero (2016) showed that the decision to lie vs. tell the truth was associated with expected consequences of stating the truth, but not with expected consequences of deceiving. Cassidy, Wyman, Talwar, and Akehurst (2019) studied the relationships between Walczyk's model variables (expectations from lying and truth telling) and decisions to lie vs. tell the truth varying motives for lying (either benefit to oneself or another person) and potential costs (either to oneself or another person). Contrary to Masip et al. (2016), they found no relation between the expected value of truth telling and decision to lie in self-oriented lies,[1] whether or not there was a cost to another, whereas they found a significant negative relationships between the expected value of truth telling and decision to lie, when the lie was other-oriented. Overall, when there was an additional cost of lying — implicating the other person of wrongdoing that self had conducted, or implicating self for wrongdoing the other had conducted — the participants were less prone to lie.

Building on this previous research, using the data from the scenario-based study by Sakamoto et al. (2013), which varied riskiness, domains of costs and benefits, severity of loss when detected, and motives for deceiving, we wanted to test if ADCAT model would make the same predictions as their results suggested, namely that the potential outcomes from (successfully or unsuccessfully) deceiving do matter in decisions to deceive. Indeed, we found that in most conditions the expected value of lying both correlated with and predicted decisions to lie better than the expected value of telling the truth.

The rest of the paper is organized as follows. We start by briefly discussing Walczyk's ADCAT model and its empirical test by Masip et al. (2016). Then we review the experimental settings and the data of the Sakamoto et al. (2013) study. We continue by describing our modeling approach and presenting the results. We conclude with a discussion of possible reasons why our findings differ from those of Masip et al. (2016).

## The ADCAT model

The *Activation-Decision-Construction-Action Theory (ADCAT)* by Walczyk et al. (2014) is a cognitive model of deception (with high stakes) specifying the roles of cognitive, emotional, motivational, and social processes in the decisions to deceive, explicitly accounting for constructs such as working memory and theory of mind. The theory is more elaborate than is required for the current purposes, so it is presented here only to the extent that it applies to "quasi-rational" (Walczyk's term) decision making in deception context, omitting for instance discussion on cognitive load or lie construction.

### Quasi-rational decision component

According to the model the decision to deceive is influenced by both the emotional reaction to the choice options and the social context, and the decision maker chooses an option that best achieves her goal given her estimates of utilities and likelihoods of the outcomes, i.e., expected costs and benefits of choosing a particular option. She estimates the expected value of action $a$ (either lying or telling the truth) based on action's $n_a$ different possible outcomes using the equation:

$$EV(a) = \sum_{i=1}^{n_a} p_i(a) v_i(a),$$

where $v_i(a)$ is the value of $i$th outcome (gain or loss) of action $a$, and $p_i(a)$ its probability. She then chooses the action with the highest expected value. Finally, she uses these estimates to assess an overall level of motivation to lie, $M$. This is done as if she were intuitively following the equation:

$$M = EV_{lying} - EV_{truthtelling}.$$

The higher the value of $M$, the more likely she will lie, and spend cognitive resources in constructing the lie. Basically, $M$ determines which particular lie she will tell; for instance, whether she chooses to fabricate a story rather than just omit a crucial piece of information, the latter consuming less resources (Walczyk et al., 2014).

Walczyk et al. (2014) emphasizes that cognitive processes underlying deception and truth telling do not differ, but memory processes, decision making, and problem solving are essential in both. Particularly, in lying, the truth, goals, and the social context are activated in the working memory, which in turn bring in relevant life memories of previous decisions, which then control the motivation for dishonesty, while means-end problem solving is used by the liar to move from the current state to the desired goal state.

### Masip et al. (2016)'s study

To empirically test the ADCAT model's decision component, Masip et al. (2016) administered two separate questionnaires. In Questionnaire 1, the participants read ten scenarios and made binary choices between lying and telling the truth in those scenarios. The authors correlated these choices with the expected values $M$, $EV_{lying}$, and $EV_{truthtelling}$ calculated from the participants' responses in Questionnaire 2, in which they read again the same ten scenarios, and for each scenario generated a possible consequence of telling the truth, probability of that consequence, and how good or bad it would be. They were also asked to come up with an alternative consequence of telling the truth and indicate its valence. Finally, they were asked to think about what kind of lie they would tell to avoid the negative consequences of disclosing the truth, and how likely they expected it to go undetected, and the consequence of not getting detected.

They classified the participants into liars and truth tellers based on the expected values, and tested how well the classification matched the actual choices in Questionnaire 1. Their results showed that for both actual liars and truth tellers (in Questionnaire 1) the expected value of telling the truth was a better predictor of their choices than the expected value of lying, and it was even slightly better predictor than the motivation score $M$. From this they concluded that the expected outcome of successfully or unsuccessfully lying may not play a role in decisions to lie but what matters are the consequences of revealing the truth.

## Current experiment

Based on their findings, Masip et al. (2016) suggested that Walczyk's model could be made more parsimonious by omit-

[1]Masip et al. (2016) used mostly self-oriented lies in their scenarios.

ting the expected value of lying from the equation, and equating motivation to lie with expected value of telling the truth. They supported their argument with Levine et al. (2010)'s *veracity principle*, namely that people usually tell the truth unless it interferes with their goal attainment, and studies on pure lie aversion (Gneezy et al., 2013; López-Pérez & Spiegelman, 2012). According to the former, lying requires justification whereas telling the truth does not, and according to the latter, the act of lying has a cost regardless of its consequences. However, Sakamoto et al. (2013) showed that the perceived benefit of successfully deceiving predicted the deceptive choices, but only in loss-facing scenarios.[2] On the other hand, tax payers have been found to refrain from cheating in their taxes for fear of being audited and penalized (Alm, 2012; Slemrod, 2007). Using another dataset we wanted to study if Walczyk's model will support Masip et al. (2016)'s conclusion or if it would replicate findings of Sakamoto et al. (2013), deeming Masip *et al*'s suggestion of dismissing the expected value of lying premature.

## Method

**Participants** In the online study conducted by Sakamoto et al. (2013) on Amazon Mechanical Turk 492 participants (276 men, 214 women, 2 unknown, median age 29, age range 18-77 years) read a single scenario of a common everyday life situation, and answered several questions pertaining to the scenario. They were also asked demographic information including age, income, and education.

**Data** Data comprises of participants' responses to eight questions about potential communicative messages — either deceptive or truthful — that could be exchanged in the scenario they read. The participants were asked to imagine themselves as the protagonist in the scenario situation, and indicate the likelihood with which they would choose the deceptive message over the honest one, and to evaluate several aspects of the scenario. They gave responses on a continuous Likert type scale with only the end points labeled (e.g., very unlikely - very likely, very bad - very good): for questions Q1, and Q3 - Q8 the scale ranged from -5 to 5, and for question Q2 from 0 to 1. The questions were (with simplified wording): Q1. How likely would you lie rather than tell the truth in this situation? Q2. How likely would your lie be detected? Q3. If it was detected, how good or bad would the result be for you? Q4. If it was not detected, how good or bad would the result be for you? Q5. If you told the truth, how good or bad would the result be for you? Q6. How would telling the truth make you feel? Q7. How would lying make you feel? Q8. How truthful is the deceptive message in this situation? The question Q1 was asked first and then the remaining questions Q2 - Q8 were presented in a random order.

**Material** The study used both gain and loss facing scenarios depicting a situation in which the speaker had an incentive to deceive the hearer or hearers. The two loss conditions varied the magnitude of loss the speaker could expect to incur if getting detected (large vs. small). This was done by adding some extra context to scenarios. Example scenarios are shown in Figure 1.

In both gain and loss facing scenarios a deceptive communication option was paired with an honest option. In gain-facing scenarios — situations in which making the hearer believe something false benefited either the speaker or hearer — the deceptive choice was associated with probabilistic gain, whereas the honest choice was associated with certain no-gain. In contrast, in loss-facing scenarios — situations in which telling the truth harmed either the speaker, hearer, or someone else — the honest choice was associated with certain loss, whereas the deceptive choice was associated with probabilistic no-loss.

In addition to asymmetric outcome structures, there were other differences; in gain scenarios the non-deceptive message was formulated as an act of omission (i.e., "say nothing" as no response was required from the speaker), but in loss scenarios as an act of commission (i.e., as a truthful message that was sufficiently informative, clear, and appropriate (Grice, 1975)). This was necessary in order to keep the gain scenarios in the domain of pure gain, since for instance overtly stating the harmful truth would have introduced an outcome in the domain of losses.

Although both Sakamoto et al. (2013) and Masip et al. (2016) used quite similar scenarios their protocols had some important differences:

- While Masip et al. (2016) administered two separate questionnaires for the participants to decide between deceiving and telling the truth, and to evaluate the potential outcomes, Sakamoto et al. (2013) asked the participants to complete these two tasks in a single questionnaire.

- Sakamoto et al. (2013) used predefined consequences, and asked participants to evaluate their valences, in addition to estimating the detection probability. This design assumed that truth telling resulted in one certain (non-positive) outcome and telling a lie had two potential outcomes, one non-negative when going undetected (with probability $1 - p$) and one negative when detected (with probability $p$) (the payoff structure is presented in Table 1). Masip et al. (2016) in turn asked the participants to generate the possible consequences of both telling the truth and successfully lying in addition to indicating their probabilities and valences. These differences are summarized in Table 2.

- Finally, while Masip et al. (2016) asked each participant to respond to several scenarios, Sakamoto et al. (2013) presented each participant a single scenario randomly picked from a pool of 48 scenarios. The former also analyzed their data scenario-wise, whereas the latter pooled the responses within each condition into one set for the analyses.

[2] Most scenarios used by Masip et al. (2016) were loss-facing.

**Gain scenario.** Harry is in a job interview. He is not familiar with the field, and only moderately interested in it. Unsurprisingly, the job is beyond his skills. He is currently unemployed and desperately needs to find a job so he can support his family. When the interviewers offer Harry an opportunity to provide any additional information, he can't admit he would take any job he gets. If the interviewers find out that Harry deceived them, they won't hire him. Imagine yourself as Harry.

**[Honest option]** Say nothing.

**[Deceptive option]** "I think I am the right person to meet the challenges of this position, for which I am highly qualified."

**Loss scenario.** Stephanie belongs to a "popular" group in her high school. One day, she happens to chat with Nancy, who is associated with the "nerd" group. They discover that they share an interest in ancient South American cultures. Nancy invites Stephanie to come over on Saturday to see her collection of books and collectibles, and Stephanie agrees. However, the popular kids are also planning to meet on Saturday. When asked why she is not planning to join them, Stephanie is reluctant to admit that she is meeting one of the nerds. **(Low loss)** The popular group's unwritten rules allow some interaction with the nerds for things like getting help with homework, but it's still seen as not a very cool thing to do. If the other group members find out that Stephanie deceived them, Stephanie will be mildly embarrassed. **(High loss)** The popular group's unwritten rules forbid any interaction with the nerds. If the group members find out that Stephanie deceived them, she will be expelled from the group. Imagine yourself as Stephanie when one popular group member asks, "Why aren't you joining us this Saturday?"

**[Honest option]** "I'm meeting Nancy on Saturday."

**[Deceptive option]** "Guess what? We're getting a new dog on Saturday. I'm so excited!"

Figure 1: Examples of scenarios and response options.

Table 1: Payoff valences of telling the truth and lying for gains and losses in the Sakamoto et al. (2013) study.

| **Condition** | **Telling truth** | **Successful lie** | **Detected lie** |
|---|---|---|---|
| Gain | No gain (0) | Gain (+) | Loss (−) |
| Loss | Loss (−) | No loss (0) | Loss (−) |

Table 2: Outcome structure of telling the truth and lying in the studies by Masip et al. (2016) and Sakamoto et al. (2013).

| **Study** | **Telling truth** | **Lying** |
|---|---|---|
| Masip et al. (2016) | Outcome 1 (−) $p$<br>Outcome 2 (+) $1-p$ | Outcome (+) $1-p_d$ |
| Sakamoto et al.(2013) | Outcome (0/−) $p=1$ | Outcome 1 (+/0) $1-p_d$<br>Outcome 2 (−) $p_d$ |

$p$ is the probability of the outcome, and $p_d$ is the detection probability. (+/0/−) marks the valence of the outcome.

**Procedure** In the study by Sakamoto et al. (2013) the participants indicated their likelihood of lying vs. telling the truth using a continuous scale with the ends marked with the honest message at the left end and the dishonest message at the right end of the scale. This scale was interpreted such that any choice to the right of the mid-point of the scale (indifference) meant that they would more likely lie than be honest, and the choice at the extreme right end of the scale meant that they would definitely lie (i.e., with 100% probability). For the current analysis purposes these continuous choices were discretized so that values above zero were coded as 1 (lie), and values at zero and below, were marked as 0 (tell truth).[3]

We calculated the expected values of lying and telling the truth using the utilities and probabilities of the outcomes (assessed in questions Q2-Q5) with the following equations:

$$\begin{aligned} EV_{lying} &= (1-p_d)v_{no\_d}(lie) + p_d v_d(lie), \\ EV_{truthtelling} &= v(truth), \end{aligned}$$

where $p_d$ is the detection probability, and $v_d()$ and $v_{no\_d}()$ values of outcomes when detected and not detected. For the sake of simplicity, like Walczyk et al. (2014) and Masip et al. (2016), we assumed that any affective reactions (e.g., guilty feeling) were included in the outcome utility, instead of incorporating them explicitly in the equations (with values obtained from questions Q6 and Q7).

In order to test how well each of the three expected values could tell apart liars from truth tellers (i.e., predict the actual lying decisions from the expected values and $M$) we first dichotomized these values following the methodology used by Masip et al. (2016). Briefly, we first calculated the proportion (say $X\%$) of participants who chose to lie based on earlier discretized choice values. In our predictions, we

[3]Basically, we coded as liars those participants who indicated that they will more likely lie than tell the truth.

then matched this true proportion by predicting $X\%$ of participants with lowest $EV_{truthtelling}$ (and with highest $EV_{lying}$ and $M$) to be liars. Finally, in each condition, we compared these predictions to the actual choices, and calculated how many predictions were correct. Like Masip, we also computed this prediction accuracy separately for actual liars and truth tellers.

## Results

For the data analysis we pooled all participants' responses together by the three scenario conditions: gain (N=161), large loss (N=162), and small loss (N=169).[4]

We started by correlating the participants' decisions to deceive with the motivation to deceive, and expected values of successfully deceiving and telling the truth. These correlations are shown in Table 3. While expected value of truth telling had the weakest correlation with the decisions to lie, $M$ had the strongest in all three conditions, expected value of lying being in between.

Table 3: Correlations between decisions to lie and the expected values calculated from ratings in questions Q2-Q5.

| Condition | $EV_{truthtelling}$ | $EV_{lying}$ | $M$ |
|---|---|---|---|
| Gain | -0.3760 | 0.4684 | **0.5360** |
| Large loss | -0.3127 | 0.3268 | **0.4128** |
| Small loss | -0.3271 | 0.5236 | **0.5733** |

The percentages of correctly identified liars and truth tellers, using the dichotomized expected values as described above, are shown in Table 4. Just like Masip et al. (2016), we did not achieve perfect identification: all individuals who indicated that they would lie (tell the truth) were not coded as liars (truth tellers) by their dichotomized expected values. However, even if the scenarios and conditions were not exactly similar in these two studies, the overall identification accuracies were quite close: the identification rates averaged over their four scenarios retained for analyses were 67.11%, 62.40%, and 68.07%, for $M$, $EV_{lying}$, and $EV_{truthtelling}$, respectively, whereas our rates, averaged over the three conditions in Sakamoto et al. (2013)'s data, were 75.59%, 69.83%, and 66.06%, respectively.

For both the actual liars and truth tellers, and overall, the motivation to lie was the best predictor of their choices in the conditions involving gains and large losses. For small losses, the expected value of lying was the best predictor. While the motivation to lie predicted the choices best in large losses, that was the only condition in which the expected value of truth telling was more accurate than the expected value of lying, although the differences were not large.

Interestingly, while regression analysis run by Sakamoto et al. (2013) indicated that the outcome of successful deception

[4]We could have analyzed the data per given scenario, but this would have resulted quite small sample sizes, with about ten datapoints per scenario.

Table 4: Correct identification rates of liars and truth tellers.

| Predictor | Condition | Percentile for cutoff | Identification % Liars | Truth tellers | Overall |
|---|---|---|---|---|---|
| $M$ | Gain | $P_{50}$ = -0.37 | **77.78** | **77.50** | **77.64** |
| | Large loss | $P_{51}$ = 0.02 | **68.75** | **69.51** | **69.14** |
| | Small loss | $P_{60}$ = 0.49 | 67.16 | 78.43 | 74.00 |
| $EV_{lying}$ | Gain | $P_{50}$ = -0.38 | 71.60 | 71.25 | 71.43 |
| | Large loss | $P_{51}$ = -0.96 | 61.25 | 62.20 | 61.73 |
| | Small loss | $P_{60}$ = -0.47 | **70.15** | **80.39** | **76.33** |
| $EV_{truth}$ | Gain | $P_{50}$ = 0 | 63.00 | 73.75 | 68.32 |
| | Large loss | $P_{49}$ = -1.25 | 62.50 | 63.41 | 63.00 |
| | Small loss | $P_{40}$ = -1.54 | 56.72 | 73.53 | 66.86 |

Table 5: Comparison of logistic regression models

| Condition | $\beta_L$ | $\beta_T$ | Δ BIC $EV_L$ | $EV_T$ | $M$ | LLR p-values $EV_L$ | $EV_T$ | $M$ |
|---|---|---|---|---|---|---|---|---|
| Gain | 1.9 | -1.8 | 13 | 26 | -5 | $2\times10^{-5}$ | $3\times10^{-8}$ | 0.8 |
| Large loss | 1.3 | -1.1 | 8 | 12 | -5 | $2\times10^{-4}$ | $4\times10^{-5}$ | 0.6 |
| Small loss | 2.9 | -1.6 | 15 | 62 | 4 | $7\times10^{-6}$ | $3\times10^{-16}$ | 0.003 |

predicted the decisions to lie in losses, but not in gains, our results indicated the opposite; the expected value of deception — incorporating outcomes of both successful and failed deception — predicted the decisions to lie and tell the truth more accurately in gains than in losses. Also the expected value of truth telling was slightly better predictor of both decisions in gains than in losses.

Since $M$ is the difference between $EV_{lying}$ and $EV_{truthtelling}$, one might argue that it constitutes a more complex model. We therefore also conducted logistic regression analyses and used the Bayesian Information Criterion (BIC) to compare the evidence for a model with two predictors,

$$P_{\beta_L,\beta_T}(lying) \propto e^{\beta_L EV_L + \beta_T EV_T},$$

against the models using only one of these predictors, i.e., models in which one of the β-coefficients was forced to be zero, or in case of $M$, forcing $\beta_T = -\beta_L$. Models being nested, we also computed the statistical significance of the difference of log-likelihoods of the models.

Our model selection analysis (Table 5) clearly favors using both the value of lying and the value of truth telling in explaining decisions to lie. The BIC difference (Δ BIC) greater than 6 is generally seen as a strong evidence against simpler hypothesis and the difference more than 10 as very strong, while negative values of BIC favor simpler hypotheses. The likelihood ratio test (LLR) for comparing logistic regression models show that the differences in model fits are statistically very significant (see Table 5).

The β coefficients in fitted two-predictor models also automatically recovered the structure highlighting the role of the difference $EV_{lying} - EV_{truthtelling}$. In the gain and large loss conditions these two predictors are almost equally weighted, while in the small loss condition the expected value of lying is twice as important as the value of telling the truth. This is also highlighted in Table 5 where the negative $\Delta$ BIC suggests using $M$ instead of $P_{\beta_L,\beta_T}$ in the gain and large loss conditions. The small loss condition is better modeled using separate weights for $EV_{lying}$ and $EV_{truthtelling}$.

Our findings are somewhat contrary to what Masip et al. (2016) found, and it may be because the negative outcomes from truth telling that we gave our participants were not judged severe enough to lie, but instead the lying behavior was driven by the expected positive outcome from successfully deceiving. In turn, Masip et al. (2016)'s participants either may not have been optimistic about their lie succeeding and being helpful, or did not find the outcome from lying (that they themselves gave in Questionnaire 2) attractive enough, and therefore were driven by the very bad outcome from telling the truth, which they wanted to avoid.

## Discussion

Decisions to deceive may be driven by two "opposite" motives: an attempt to avoid a loss from harmful truth or an attempt to gain something by lying. These two perspectives may explain the asymmetries between the two studies by Masip et al. (2016) and us. They started with the consequences of being honest by asking what are the possible outcomes if the harmful truth is revealed, contrasting them to potential consequences of successfully deceiving. In contrast, Sakamoto et al. (2013) focused on the risky aspect of decisions to deceive, and asked participants to evaluate benefits of successfully deceiving and costs of getting caught, while assuming that telling the truth only had bad consequences. In fact, in their scenarios truth telling and detection shared the same outcome (e.g., not being hired), which the decision maker tried to avoid by lying. In both cases the truth got revealed, which resulted in either a loss (loss scenarios) or no gain (gain scenarios), but in case of detection there was an additional (implicit) cost of being stigmatized as a liar. In turn in gain scenarios status quo always persisted (e.g., the hearer still felt bad about her looks), but the liar incurred a cost of getting caught.

Furthermore, since Masip et al. (2016) asked their participants to focus on truth and its consequences, they may have judged its harmfulness more severely than benefits of undetected deception. In turn, Sakamoto et al. (2013) gave the outcomes and their valences to their participants, focusing on the benefits of deception as opposed to cost of telling the truth, which may have lead the participants to judge benefit of successful deception more extremely than harmfulness of truth (or failed deception). This could be the reason why in our study the motivation to lie (difference between value of lying and value of telling the truth) predicted participants' choices the most accurately. While both studies contrasted the outcome from telling the truth (bad or very bad) to successful deception, in Masip et al. (2016)'s study the outcome from undetected deception could still have been somewhat bad, whereas in Sakamoto et al. (2013)'s study it was assumed to be very good. Finally, it seems that it also mattered if by deceiving the participants could gain something or avoid losing something, and that determined whether they focused on benefits of lying or costs of telling the truth, in other words, the reference point they adopted, as suggested by the prospect theory (Kahneman & Tversky, 1979).

## References


Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*, 96–104.

Alm, J. (2012). Designing alternative strategies to reduce tax evasion. In A. P. Michael Pickhardt (Ed.), *Tax evasion and the shadow economy* (pp. 13–32). Edwar Elgar Publishing.

Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2013). When do we lie? *Journal of Economic Behavior & Organization*, *93*, 258–265.

Cassidy, H., Wyman, J., Talwar, V., & Akehurst, L. (2019). Exploring the decision component of the activation-decision-construction-action theory for different reasons to deceive. *Legal and Criminological Psychology*, *24*(1), 87-102.

Coricelli, G., Rusconi, E., & Villeval, M. C. (2014). Tax evasion and emotions: An empirical test of re-integrative shaming theory. *Journal of Economic Psychology*, *40*, 49–61.

dePaulo, B. M., Kashy, D. A., Kirkendol, S. E., & Wyer, M. M. (1996). Lying in everyday life. *Journal of Personality and Social psychology*, *70*(5), 979–995.

Dhami, S., & al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior & Organization*, *64*, 171–192.

Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, *58*, 723–733.

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization*, *93*, 293–300.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics volume 3: Speech acts* (pp. 41–58). New York: Academic Press.

Gupta, S., Sakamoto, K., & Ortony, A. (2013). Telling it like it isn't: A comprehensive approach to analyzing verbal deception. In F. Paglieri, L. Tummolini, R. Falcone, & M. Miceli (Eds.), *The goals of cognition. Essays in honor of Cristiano Castelfranchi.* London: College Publications.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Levine, T. R., Kim, R. K., & Hamel, L. M. (2010). People lie for a reason: Three experiments documenting the principle of veracity. *Communication Research Reports*, *27*(4), 271–285.

López-Pérez, R., & Spiegelman, E. (2012). Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics*, *16*(3), 233–247.

Maciejovsky, B., Schwarzenberger, H., & Kirchler, E. (2012). Rationality versus emotion: The case of tax ethics and compliance. *Journal of Business Ethics*, *109*, 339–350.

Masip, J., Blandón-Gitlin, I., de la Riva, C., & Herrero, C. (2016). An empirical test of the decision to lie component of the Activation-Decision-Construction-Action Theory (ADCAT). *Acta Psychologica*, *169*, 45–55.

Sakamoto, K., Laine, T., & Farber, I. (2013). Deciding whether to deceive: Determinants of the choice between deceptive and honest communication. *Journal of Economic Behavior & Organization*, *93*, 392–399.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, *21*(1), 25–48.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology*, *34*, 22–36.