Gaze signatures of cognitive conflict while choosing and solving

Revati Shivnekar (revati.shivnekar@uj.edu.pl)
INCET, Jagiellonian University, Krakow 31-044 Poland
Department of Cognitive Science, Indian Institute of Technology, Kanpur 208016 India

Nisheeth Srivastava (nsrivast@iitk.ac.in)
Department of Cognitive Science, Indian Institute of Technology, Kanpur 208016 India

Abstract

Current measures of cognitive conflict in experimental settings focus either on whole trial-level measures, such as reaction time and proportion of cohort disagreement, or intrusive on-task measures such as think-aloud paradigms. Consequently, granular within-trial measurements of the experience of cognitive conflict have been missing from the literature, and consequently, from formal models and theories of decision making. By combining the recently proposed switch paradigm for measuring cognitive conflict with on-task eye-tracking, we ask one such theoretical question: is the experience of cognitive conflict different when choices have clear normative answers and when they don't? Our results answer this question affirmatively and characterize it quantitatively by means of gaze signatures for both classes of experience of cognitive conflict.

Keywords: conflict; decision making, eye tracking

Introduction

Rules and heuristics often simplify decision-making. Constitutions, for instance, outline rights and regulations that aim to guide behavior, while systematic methods, such as a method for solving equations, can be broadly applied to specific problems. Strategies are said to make decision-making standards explicit, offering a level of abstraction that ensures generalizability and providing a framework for evaluating outcomes (Levine, Chater, Tenenbaum, & Cushman, 2023).

However, not all decisions align with strategic problem solving. Many aspects of daily life resist such approaches due to conflicting motivations and lack of objective bases for decisions (e.g., should a doctor perform a legal abortion despite personal ethical objections? What treatment should I choose for an ailing parent?). In such scenarios, the experience of conflict can be phenomenologically prolonged, especially when strategies are unknown or their application is ambiguous. This paper explores cognitive conflict as a dynamic experience, examining its temporal evolution in reasoning tasks. Specifically, it contrasts choice contexts where the conflict is expected to be more sustained due to the absence or ambiguity about choice resolution strategy with those where it is expected to be constrained to specific periods due to the existence of such clearly applicable strategies.

Temporal dynamics of conflict

Empirical studies tend to describe cognitive conflict fairly mechanically as cognitive processes competing to support different alternatives (e.g., Evans (2007); Greene et al. (2009)) or the simultaneous activation of incompatible mental representations (e.g., Botvinick, Braver, Barch, Carter, and Cohen (2001)). Measures from such investigations often distill this rich and dynamic experience into singular metrics, such as reaction times or subjective ratings, overlooking the changes in preferences that are integral to this experience. Conflict is a keenly conscious phenomenon experienced in time, particularly in contexts involving reasoning and problem solving. Our thoughts seem to unfold sequentially, often leading us to mentally switch preferences as different arguments present themselves. Similarly, some have proposed that decision-making in complex tasks can be modeled like perceptual decision processes, where choices emerge through the stochastic accumulation of evidence. In this view, conflict emerges dynamically as certain alternatives temporarily dominate, only to be supplanted by others, often in a seemingly random fashion (Gürçay & Baron, 2017; Krajbich, Lu, Camerer, & Rangel, 2012; Pärnamets et al., 2015).

This fluidity of the decision-making process is amplified when making decisions without a clear strategy or systematic framework to compare the aspects of alternatives to reach a conclusion. Consider, for example, the footbridge version of the trolley problem, a runaway trolley is set to kill five workers unless intercepted by pushing a heavy person onto the tracks. Initially, the goal of saving more lives may lead me to consider pushing, but uncertainty about my ability to push the person or the possibility of him fighting back may shift my preference toward not intervening. As we reason, different aspects of the problem can momentarily influence our preferences. There may also be no fixed temporal order in how these preferences arise (alternatively, see Greene, Sommerville, Nystrom, Darley, and Cohen (2001)). As a result, in the absence of a clear way to compare these different aspects and decide, the conflict between alternatives may persist throughout the deliberation, until we either settle on a choice or make a decision due to external task demands.

Contrast the phenomenological complexity of confronting a moral dilemma such as trolley problems with the experience of possible cognitive conflict while solving puzzles such as logical syllogisms. A syllogism consists of two premises that are assumed to be true and a conclusion. The task is to determine whether the conclusion logically follows from the premises. Syllogisms can either be valid or invalid, and their validity can be assessed using learned strategies like Venn di-

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

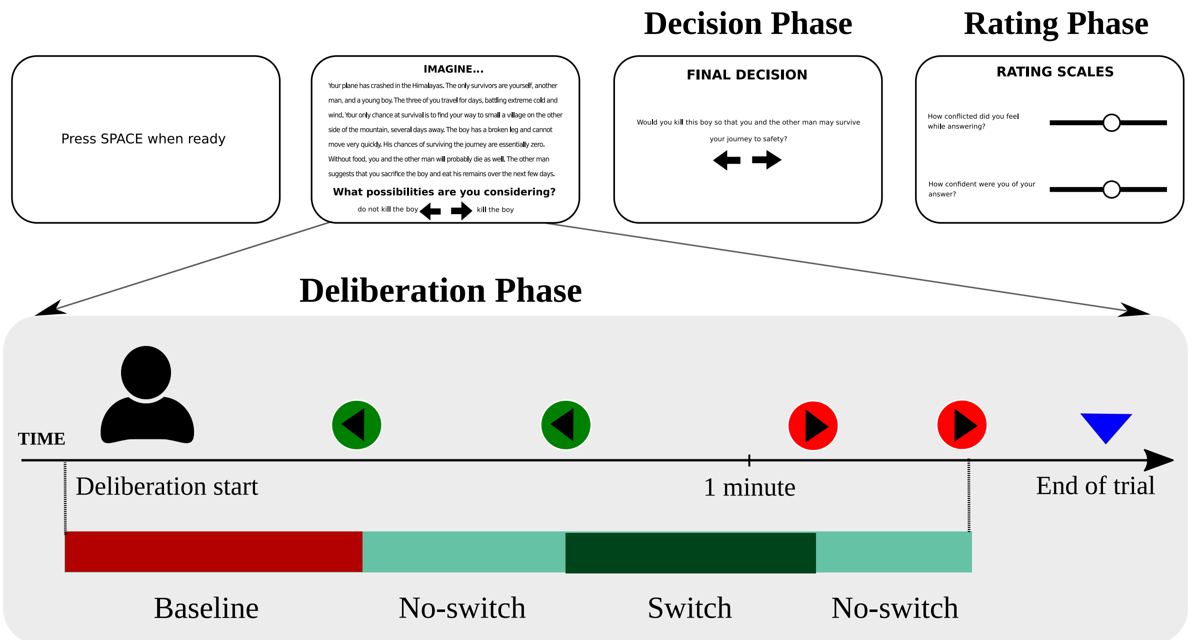

Figure 1: Schematic of the switch method and interest periods defined on the gaze data for Experiments 1 and 2. The top section illustrates a typical trial structure with deliberation, decision and rating phases. The inset shows the deliberation phase in a representative trial. In this trial, four key presses indicating interim preferences of the participant are recorded during the deliberation phase (green and red circles) until the trial ended (blue triangle). These key presses were used to segment gaze data into blocks: baseline (data collected before any preference was recorded), no-switch (consistent consecutive key presses), and switch (changes in consecutive key presses).

agrams or truth tables. For example, in a type of syllogisms which are typically difficult to solve, called multiple-model syllogisms, the premises can be combined in more than one way. The challenge here is to decide if the conclusion is valid in *all* such possible arrangements. One way to solve these syllogisms is to look at one arrangement at a time and check if the conclusion follows. If a series of valid arrangements confirms the conclusion, each arrangement will only add to the certainty of the syllogisms validity. However, if one arrangement does not match the conclusion, the reasoner may feel conflicted, wondering if the syllogism is truly invalid or if they applied the strategy incorrectly. Hence, conflict in solving these difficult syllogisms would be limited to the periods when participants reach contradicting judgments consecutively.

In this paper, we empirically explore how the experience of conflict may be more prolonged when choosing between alternatives in a moral dilemma compared to solving a syllogism when one is familiar with and trained in the method. People do not always prefer the outcome that maximizes the number of lives saved, for example, as preferences are often influenced by contextual factors such as the action itself (Greene et al., 2001), the mechanism behind it (Greene et al., 2009), the framing effects (Rehren & Sinnott-Armstrong, 2021), etc. Our own experiences of deliberating on difficult ethical issues suggest that we rarely have a strict preference for one ethical principle. As we recognize the pull of various arguments, we hypothesize that participants will experience conflict throughout their engagement in such dilemmas. In contrast, when solving syllogisms using learned techniques, the conflict is more likely to arise only when a dissimilar answer is reached after solving the same syllogism. Thus, the conflict will be limited to moments in reasoning when the answer does not match the one previously reached.

We used the *switch* paradigm to detect changes in momentary preferences (Shivnekar & Srivastava, 2024). Participants were presented with either a moral dilemma or a syllogism, each offering two alternatives, identified by the left or right arrow keys. The inset in Figure 1 illustrates a typical deliberation phase during a trial. Participants indicated their leading choice throughout the reasoning process by pressing the corresponding keys, revealing periods when they stuck with their previously reported choice (no-switch block) or switched to a different option (switch block). We hypothesized that trials in which participants switched between alternatives more frequently would be perceived as more conflicting.

To track the temporal dynamics of conflict with high granularity, we used eye tracking, specifically focusing on pupil dilation and fixation durations as indicators of conflict. While pupil dilation has been used as an indicator of cognitive effort and exploration of strategies in reasoning tasks before (see Wang and Chen (2019) for a review), we also examined fixation durations as a measure of conflict, hypothesizing that longer fixations reflect increased attention to problem-relevant details during decision-making (Pieters & Warlop, 1999).

Experiment 1

Method

Participants Twenty-seven participants were recruited through email advertisements. We had to exclude two participants from the analyses as their data files were corrupted before analyzing the data. Final analyses were performed on data from 25 participants (3 females; Mean age = 22 years).

Experiment Design and Procedure The participants solved 16 problems which were either of the type non-moral or moral. A moral dilemma was either low-conflict personal (LC), high-conflict personal (HC), or impersonal (IM). Each problem - moral or non-moral - was written in the second person and gave the participant a choice between an action and its omission. Non-moral problems were expected not to invoke any moral principles. Actions in personal moral dilemmas (LC, HC) were of the sacrificial type with a choice between a utilitarian action (henceforth, U) that saved many by killing few and its omission (henceforth, D for deontological). The actions in IM dilemmas did not involve physical proximity but were moral transgressions such as bribing, stealing, etc. All stimuli were obtained from Koenigs et al. (2007), who classified the problems as low or high conflict based on the degree of disagreement observed between the participants in their study, a categorization that was broadly supported by our previous investigation (Shivnekar and Srivastava (2024)).

Participants reported their interim preferences within the switch paradigm. The trials were self-paced with three phases: deliberation, decision, and rating. During the deliberation phase, participants read a problem and responded to the following question: ‘What possibilities are you considering?’ In moral problems, the D and U choices were linked to the left and right arrow keys, respectively, with each choice explicitly shown on the screen with arrow images. Participants were instructed to monitor their thoughts and press the corresponding key to indicate any momentary preference they wished to report during this phase by pressing the associated key. Multiple responses were allowed, but trials without key presses during this phase were excluded. The deliberation phase lasted a minimum of one minute, but could extend as long as the participant needed to reason. Participants recorded their final judgment in the decision phase. Lastly, they rated how conflicted they felt while deliberating and their confidence in the final decision they made in the rating phase of the trial. These ratings served as validity checks for the vacillations method.

We tracked participants’ eye movements monocularly while they deliberated on a problem using Eyelink 1000 Plus (SR Research Ltd.) at a sampling rate of 1000 Hz. To ensure stability, the participants’ heads were supported with a

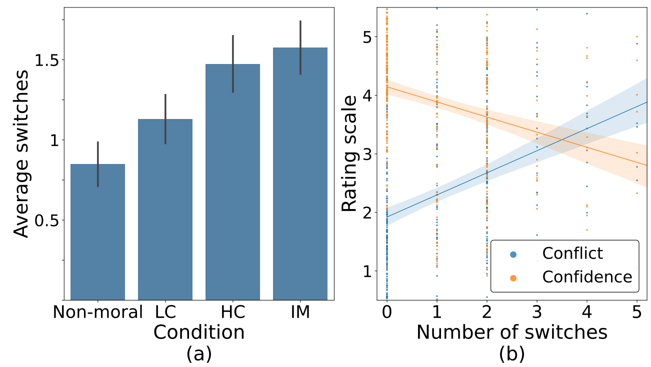

Figure 2: (a) Average switches in each condition with standard error bars and (b) jittered correlation plot of ratings (conflict and confidence) by number of switches in Experiment 1.

chin rest and were seated 65 cm away from a 1920 x 1080 pixel display. A nine-point calibration and validation procedure was conducted at the start of the experiment to ensure accurate tracking. The drift correction was assessed before each trial, with a re-calibration performed as needed to maintain precision.

Preprocessing of gaze data To investigate gaze patterns in reasoning, we defined the interest period for the deliberation phase alone, beginning when participants were first presented with the problem and ending when the last choice key was pressed within the deliberation phase. Data recorded after the final key press that indicated a preference during the deliberation phase and before the trial ended were excluded, as we could not be sure if the participants were still engaged in reasoning or simply waiting for the mandatory deliberation period to end. We then segmented the interest period into blocks based on preference changes indicated by consecutive key presses. These blocks captured how the participants reengaged with the problem. Data between consecutive identical key presses (right-right, left-left) were categorized as no-switch blocks, indicating no change in preference. In contrast, data between dissimilar consecutive key presses (right-left, left-right) were categorized as switch blocks, reflecting preference changes during reasoning. Data preceding the first key press—during which participants were reading the problem for the first time and also likely reasoning about it concurrently—were treated as a separate block (see Figure 1 for a schematic). This block also served as a baseline to compare how participants reengaged with the problem after reading it and communicating the first preference. Key presses occurring within 1500 ms of one another were excluded to avoid excessive fragmentation of the data.

Fixation durations were smoothed using a rolling average with a window of 10 observations, excluding missing values. The size of the pupils was measured in arbitrary units during fixations and normalized per participant to allow comparisons. These normalized data also allowed us to infer pupil constriction or dilation relative to average pupil size (for a similar approach, see Purcell, Roberts, Handley, and Howarth (2023)). For pupil dilation analyses, we removed samples recorded off-screen, around blinks (100 ms before and after), and isolated erroneous samples. Missing pupil data were interpolated with cubic splines and down-sampled to 25 Hz.

Results

More vacillations, more conflict As anticipated, the participants reached a more unanimous consensus in rejecting the action in the LC dilemmas compared to the HC and IM dilemmas, consistent with Koenigs et al. (2007) (mean endorsement rates with 95% confidence intervals: *LC = .15 [.08, .22], HC = .66 [.57, .75], IM = .48 [.38, .58]*). Even then, across the three moral categories, participants frequently vacillated, switching between alternatives during deliberation. In particular, trials with more frequent vacillations were also perceived as more conflicting, with each switch in preference adding 0.29 to the conflict rating (Bayesian hierarchical model of conflict ratings with participant and trial as random effect with [CrI][1]: *Intercept = 2.19 [1.89, 2.49], Switches = 0.29 [0.22, 0.35], SD(Trial) = 0.65 [0.39, 1.06], SD(Participant) = 0.30 [0.09, 0.52]*). The confidence of the participant in the final reported judgment was also lowered by 0.22 on the rating scale with each added switch on the preference (Confidence model [CrI]: *Intercept = 3.87 [3.60, 4.14], switches = -0.22 [-0.28, -0.15], SD (Trial) = 0.54 [0.33, 0.90], SD(Participant) = 0.30 [0.11, 0.49]*). Refer to Figure 2 for an overview of switches between categories and subjective ratings based on the number of switches.

We also inspected the order in which preferences shift in a trial. DU and UD transitions—indicating that the first and last key presses were different—occurred with equal frequency. In trials where the first and last choices were both D or both U, at least 39% and 33% of cases, respectively, involved two or more switches between the same two keys. This indicates frequent, unordered vacillations, challenging predictions made by models such as the classical dual-system model (see Greene et al. (2009)).

Signs of sustained conflict during re-engaged moral reasoning Before any key is pressed during deliberation to report an interim preference, participants are likely to think along with reading the dilemma. Since this period was long enough, it served as a reasonable baseline period.

We compared fixation duration and pupil size in the baseline condition to the no-switch and switch blocks (Figure 3). The fixations of the participants were longer than the baseline by close to 39 ms and 57 ms in the non-switch and switch blocks, respectively (Estimates [CrI]: *Intercept = 231.84 [231.03, 232.65]; Non-switch = 38.81 [37.30, 40.34]; Switch = 56.86 [54.79, 58.96]*). The pupils were also dilated in both blocks compared to the baseline (Estimates [CrI]: *Intercept = -0.09 [-0.09, -0.08]; No switch = 0.15 [0.15, 0.16]; Switch = 0.27 [0.26, 0.28]*). Hence, fixation duration were longer and pupil size dilated regardless of the block compared to the baseline. These results, although exploratory, suggest that when reasoning about moral dilemmas, conflict remains sustained even after an initial hunch, likely due to lack of a clear preference for an alternative or shifting focus on different aspects of the problem. As a result, the reasoning process in no-switch blocks resembles that of periods when preferences shift.

[1] All Bayesian models were run in 4 chains and 20000 iterations, half of which were warmup draws with informed priors. CrI is the 95% credible interval.

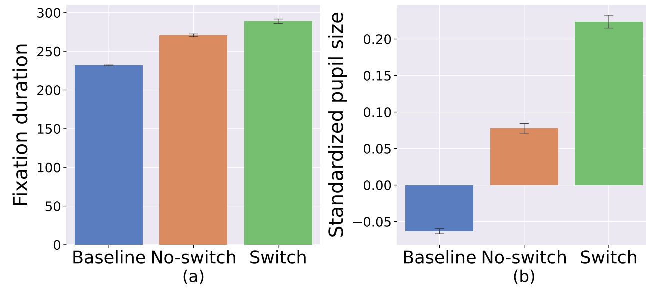

Figure 3: Blockwise mean and bootstrapped 95% CIs for (a) fixation duration and (b) standardized pupil size by preference blocks in Experiment 1.

Experiment 2

Next, we compare these findings to the conflict patterns observed in syllogism tasks.

Method

Participants We expected fewer switches this time, since most participants were familiar with solving syllogisms. Hence, to roughly match the number of switch blocks in Experiment 1, we hired a larger group of participants. After excluding eight participants due to poor quality of gaze data, the final sample consisted of 62 participants (18 females; Mean age = 21.80 years).

Experiment Design, procedure, and preprocessing Participants solved 8 categorical syllogistic reasoning problems. We employed a within-subject 2x2 design with two factors: validity (whether the conclusion follows logically from the premises) and believability (whether the conclusion is believable). Our stimuli were also divided into one of the two forms, single or multiple models (Newstead, Pollard, Evans, & Allen, 1992). The single-model syllogisms (hereafter SM) were in AAA mood and were taken from Robison and Unsworth (2017). For example:

All A are B.
All B are C.
Therefore, all A are C. (valid)

Multiple-model syllogisms (MM) are those for which more than one conceptually distinct model can be constructed from different arrangements of the premises. To necessarily conclude on the validity of the conclusion, all models need to be considered. These were in IEO or EIO mood, taken from Evans, Barston, and Pollard (1983). An example of IEO syllogism is below:

Some A are B.
No B are C.
Therefore, some A are not C. (valid)

The trial structure was similar to Experiment 1. Instead of a vignette, two premises and a conclusion were displayed on separate lines during the deliberation phase. The prompt was updated to 'Which option are you considering?' for clarity. The participants used the right and left arrow keys to indicate whether the conclusion was TRUE or FALSE, respectively. They also gave subjective conflict and confidence ratings on the last screen. The same preprocessing procedure from Experiment 1 was applied in Experiment 2, dividing the fixation duration and pupil size data into baseline, no-switch and switch blocks.

Results

Belief bias and vacillations Most of the participants (55 of 62) were familiar with the syllogisms. This familiarity suited our goal of investigating eye movements in problem solvers. Although familiar, belief bias was observed in both the SM and the MM syllogism such that believable syllogisms were more likely to be accepted as valid (SM: $F(1, 61) = 8.93, p = .003$; MM: $F(1, 61) = 14.22, p < .001$). Valid syllogisms were more likely to be judged (correctly) as valid (SM: $F(1, 61) = 204.77, p < .001$; MM: $F(1, 61) = 81.76, p < .001$) with no interaction between the two factors being significant (SM: $F(1, 61) = 1.96, p = .16$; MM: $F(1, 61) = 1.04, p = .31$).

In SM syllogisms, with only one arrangement of premises, there were fewer switches compared to MM syllogisms, where premises could be arranged in multiple valid ways (one-sided paired t test: $t\ (61) = -2.6989,\ p = 0.004$). Again, the number of switches correlated with subjective ratings in the expected manner as demonstrated by the linear mixed-models with participant as the random effect. Every switch increased the average conflict rating by 0.55 points (Intercept: *Estimate (SD) = 2.03 (0.10), DF = 67.80, p < .001, SD(Participant) = .67; Switches: Estimate (SD) = 0.55 (0.07), DF = 465.94, p <* .001, *SD(Residual) = .94*) while reduced the confidence rating by 0.41 points (Intercept: *Estimate (SD) = 4.26 (0.06), DF = 73.77, p <* .001, *SD(Participant) = .1; Switches: Estimate (SD) = -0.41 (0.05), DF = 486.22, p = .003, SD (residual) = .70*).

Pupils reveal sensitivity to changes in answer while solving syllogisms In Experiment 2, we recorded participants' fixation duration and pupil data while they solved these syllogisms with the aim of investigating how participants re-visit the problem after forming and reporting an initial preference. As we have mentioned already, most participants were aware of how syllogisms are solved and had strategies to employ. Although we did not investigate specific strategies, on the whole, we expected participants to stick to the deduced conclusion until it did not fit certain arrangement of premises leading to a switch away from the currently held preference. This confirmatory approach was expected to generate conflict only before an imminent switch in response, specifically during switch blocks. In contrast, in no-switch blocks, we anticipated that participants would continue affirming their current preference, resulting in lower levels of conflict.

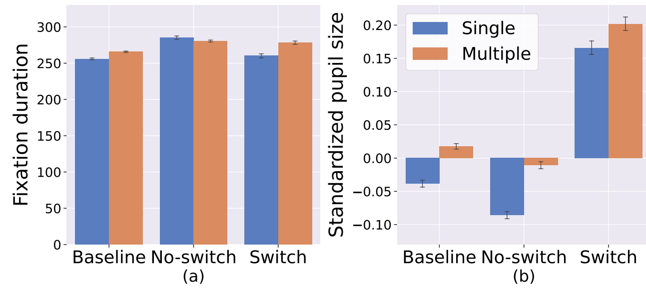

Figure 4: Blockwise mean and 95% bootstrapped confidence intervals for (a) fixation duration and (b) normalized pupil size data from Experiment 2.

Average duration of the fixations did not match our prediction. In SM, this gain was 29 ms and 4 ms in the no switch and switch blocks, respectively (Estimates [CrI]: *Intercept = 255.85 [254.09, 257.60]; No switch = 29.35 [26.88, 31.79]; Switch = 4.30 [0.35, 8.24]*). In MM, the gain was close to 15 ms and 13 ms in the no-switch and switch blocks, respectively (*Intercept = 265.70 [264.63, 266.78]; no-switch = 14.71 [13.08, 16.33]; switch = 12.56 [10.27, 14.87]*). Hence, both switch and no-switch blocks showed an increase in average fixation duration, albeit smaller, compared to the baseline.

However, pupils were more sensitive to the difference in the switch and no-switch blocks in the direction that we expected. The switch blocks recorded the largest pupil size in the SM and MM syllogisms (Figure 4 (b)). More importantly, Bayesian estimates of pupil sizes in no-switch blocks were comparable in SM (*Intercept = -0.04 [-0.04, -0.03]; No-switch = -0.05 [-0.06, -0.04]; Switch = 0.2 [0.19, 0.22]*) and constricted in MM (*Intercept = 0.02 [0.01, 0.02]; No switch = -0.03 [-0.04, -0.02]; Switch = 0.18 [0.17, 0.19]*) compared to the baseline. This notable difference between switch and no-switch blocks when reasoning about syllogisms versus moral dilemmas supports our main hypothesis.

General Discussion

Our empirical investigation described above focuses on the phenomenological experience of reasoning, contrasting problems that are amenable to strategic solutions with those that are not straightforwardly so. In solving puzzles, we typically apply a specific strategy until it fails to yield a satisfactory answer, at which point we pivot—much like changing paths in a maze after hitting a dead end, or reconfiguring premises using Venn diagrams in logical reasoning tasks. By contrast, realworld problems tend to be more complex because the standards for evaluating judgments in the wild are often neither clear nor easily deducible. For instance, in moral dilemmas determining the “correct” choice is often challenging when there is no strong preference for one ethical principle over another, with contextual nuances making the choice even harder.

To pin down the difference in choosing between alternatives in moral dilemmas and solving for a correct answer in syllogisms, we focused on the experience of conflict. We hypothesized that conflict would be sustained through time when reasoning about morally charged situations, indicating a lack of consistent preference in deliberations for any alternative. In contrast, solving logical syllogisms armed with strategies would confine conflict to periods when reasoning results in an unexpected solution.

To distinguish the tasks’ temporal signatures of conflict, it was crucial to track reasoning closely in time while minimizing intrusions from task demands. The *switch* paradigm, combined with gaze metrics, allowed us to monitor conscious preference shifts through vacillations and identify moments when the dominant preference changed. This continuous tracking offered a clearer picture of when and how preferences shifted while reasoning.

We employed vacillations as indicators of conflict. In both tasks, participants vacillated the most when solving the challenging problems (HC and MM). They also rated problems higher on conflict and lower on confidence when they had switched more frequently while deliberating. Further, vacillations allowed us to test specific predictions about the order and frequency of preference revisions from dual-process models (see Bago and De Neys (2019); Greene et al. (2001)). While default-interventionist models predict more D-to-U transitions in moral dilemmas, participants in Experiment 1 showed equal D-to-U and U-to-D shifts. Contrary to the hybrid model proposed by Bago and De Neys (2019), even participants who started and ended with the same response (DD or UU) often showed multiple vacillations before settling on a choice, suggesting that apparent consistency can mask underlying indecision.

Our main objective, however, was to test whether the experience of conflict unfolds differently over time across the two tasks. In moral dilemmas, we expected re-engagements to show a consistent pattern of conflict, regardless of whether a preference shift occurred—reflecting ongoing tension throughout deliberation. By re-engagement, we refer to instances where participants return their attention to the problem after having initially read or considered it. In contrast, for logical reasoning tasks, we anticipated that conflict would arise mainly during re-engagements that preceded a switch in judgment, indicating more localized conflict tied to the moment of change.

Pupil size estimates aligned with the predicted trend. Moral dilemmas elicited pupil dilation regardless of whether the final preference aligned with the initial one. In contrast, when solving syllogisms, participants who changed their answers exhibited greater pupil dilation, whereas those who remained consistent showed pupil constriction. Although this measure is not widely used in extended reasoning tasks due to the challenges of selecting an appropriate baseline period, previous research has shown that pupil dilation is linked with individuals exploring new strategies rather than exploit existing ones—analogous to switch blocks in our logical reasoning task (Hayes & Petrov, 2016). Our preliminary results align with these findings, suggesting that while step-by-step reasoning may confine exploration in the event of a strategy’s failing, reasoning on more complex problems may ensue exploration for an alternative throughout, possibly to a varying degree.

Our predictions for average duration of a fixation did not match the results. This measure is often used with predefined areas of interest (AoI) so as to limit the prediction to fixations on the relevant pieces of information. We did not define AoIs in our experiments due to the mandatory one-minute minimum deliberation period. Given this extended duration, it was unrealistic to expect participants to maintain consistent gaze patterns on specific AoIs, particularly for moral dilemmas where it was difficult to identify words or phrases that would be relevant to all participants. Individual differences further complicated this; some participants scanned the entire problem, while others focused on specific details.

Looking forward, future research could expand upon these findings by incorporating some design changes such as predefining AoIs, removing restrictions on deliberations etc. which will allow for a stricter interpretation of how information is processed while elucidating on gaze patterns at the individual and cohort level. A within-subjects approach will also help solidify these preliminary results. Since we used different participant pools for the two experiments, using the same participants for both tasks would allow for more direct comparisons and a clearer understanding of individual differences in conflict dynamics and preference shifts.

In conclusion, our study examined gaze signature differences during reasoning about logical and moral dilemmas, characterizing how the experience of cognitive conflict varies between these two types of tasks by correlating gaze measurements with a recently introduced within-trial conflict measure (Shivnekar & Srivastava, 2024). We show that conflict is task dependent and that its eye-tracking correlates can be tracked in real time with minimal interference. Our findings support the hypothesis that moral dilemmas may evoke sustained conflict throughout the process of deliberation, whereas logical tasks, such as syllogisms, tend to elicit conflict mainly during the reconsideration of initial choices. Secondary results demonstrate the sensitivity of pupil data in detecting cognitive conflicts and shifts within choice trials. By analyzing fixation and pupil metrics broadly, yoked to a within trial self-report paradigm for measuring cognitive conflict, we uncover new knowledge about how such conflict unfolds in time.

References

Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General*, *148*(10), 1782.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, *108*(3), 624.

Evans, J. S. B. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, *13*(4), 321–339.

Evans, J. S. B., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & cognition*, *11*(3), 295–306.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*(3), 364–371.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fmri investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Gürçay, B., & Baron, J. (2017). Challenges for the sequential two-system model of moral judgement. *Thinking & Reasoning*, *23*(1), 49–80.

Hayes, T. R., & Petrov, A. A. (2016). Pupil diameter tracks the exploration–exploitation trade-off during analogical reasoning and explains individual differences in fluid intelligence. *Journal of cognitive neuroscience*, *28*(2), 308–318.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*(7138), 908–911.

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in psychology*, *3*, 193.

Levine, S., Chater, N., Tenenbaum, J. B., & Cushman, F. (2023). Resource-rational contractualism: A triple theory of moral cognition. *Behavioral and Brain Sciences*, 1–38.

Newstead, S. E., Pollard, P., Evans, J. S. B., & Allen, J. L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, *45*(3), 257–284.

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences*, *112*(13), 4170–4175.

Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of research in Marketing*, *16*(1), 1–16.

Purcell, Z. A., Roberts, A. J., Handley, S. J., & Howarth, S. (2023). Eye movements, pupil dilation, and conflict detection in reasoning: Exploring the evidence for intuitive logic. *Cognitive Science*, *47*(6), e13293.

Rehren, P., & Sinnott-Armstrong, W. (2021). Moral framing effects within subjects. *Philosophical Psychology*, *34*(5), 611–636.

Robison, M. K., & Unsworth, N. (2017). Individual differences in working memory capacity and resistance to belief bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, *70*(8), 1471–1484.

Shivnekar, R. V., & Srivastava, N. (2024). Measuring vacillations in reasoning. *Judgment and Decision Making*, *19*, e15.

Wang, J. T.-y., & Chen, W. J. (2019). Investigating pupil dilation in decision research. In *A handbook of process tracing methods* (pp. 45–52). Routledge.