

Experimental Strategy for Investigating the Neural Basis of Framing Effects

Sacha Bourgeois-Gironde, Elise Payzan

► **To cite this version:**

Sacha Bourgeois-Gironde, Elise Payzan. Experimental Strategy for Investigating the Neural Basis of Framing Effects. Summer School of Behavioral Economics, Jul 2005, Budapest, Central European University, 2005. <ijn_00000665>

HAL Id: ijn_00000665

https://jeannicod.ccsd.cnrs.fr/ijn_00000665

Submitted on 11 Jan 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Behavioral and Neural foundations of Framing Effects

Sacha Bourgeois-Gironde¹ Élise Payzan²

¹École Normale Supérieure LSH (Lyon)
Institut Jean Nicod (Paris)

²École Normale Supérieure (Paris)
London School of Economics (London)

July 2005

- 1 Project: Pin down framing effects in situation of risky choice
 - Object of interest: framing effects for risks
 - Questions
- 2 Behavioral data
 - Robustness of this type of framing effects
 - Violation of the invariance principle and irrationality
- 3 fMRI data
 - Capture a possible feeling of discomfort while realizing ex post one's inconsistency
 - Neural underpinnings of the moral act

Risky choice type framing effects

- Subjects are asked to make a risky choice between two options that are presented in two different though logically equivalent fashion. → **Positive and negative frames.**
- Existing literature:
 - Kahneman and Tversky's famous «*Asian Disease*» problem: **risk aversion for gains - in the positive frame**, whereas **risk-seeking for losses - in the negative frame** → **Inconsistency result** achieved using a «*between-subjects*» design.
 - Sandra L Schneider (1992) fleshed out this result, using a «*within-subjects*» design, with several scenarios, not just the «*Asian-Disease*» story: **not a clear pattern for the negative frame**, while the risk aversion is confirmed for the positive frame.

Aims

- The design used by Sandra Schneider might lack of robustness, because of a **potential carryover effect**: different scenarios presented to the subjects without filler tasks.
- Test Kahneman and Tversky's inconsistency result again by using another «*within-subjects*» design.
- Hypothesis to be tested: the violation of the invariance rule is due to an extrinsic (emotional-moral) factor: **broadened rationality hypothesis**; alternative hypothesis: «**Prospect Theory**» explanation.

A multi-version of the «asian-disease problem»

- Run several trials by varying the degree of moral-emotional content of the scenarios: rate of death or live of humans, rate of failure or success of a commercial product.
- Use of a «*within-subjects*» procedure: each subject faces in turn the positive frame (henceforth PF) and the negative frame (NF) **with «filler» tasks** between frame pairs and between the different scenarios; this to try to **bypass the potential carryover effect**. ▶ Caveat
- For each problem, one risky option, versus the «sure choice».
- Look at the proportion of people making the risk averse choice in the PF and the proportion making the risk averse choice in the NF.

Predictions

- We expect «*framing effect*» to be the most frequent pattern only in the most extreme scenarios; conversely the «softer» scenarios might be associated with steady preferences across the two frames (risk aversion).
- Possible explanation: people are basically risk averse, but there is a «*moral*» dimension guiding some decisions too.
- This experimental design rests on a **strong inference** reasoning: if prospect theory [resp. the competing hypothesis] is true then we won't [resp. will] observe **sensitivity to the emotional-moral content of the scenarios**.

«Norm retrieval» test (1 out of 3)

- Zero in on the versions such that «framing effect» is the most frequent pattern - we predict «asian-disease» will be in this category; second step restricted to these cases.
- Test «the robustness» of this violation of the invariance rule, by making the latter directly accessible to the subject. → Sort of «**Revealed Preference Argument**», echoing Slovic and Tversky's «*Understanding/Accepting principle*»:
 - If framing effects stem from bounded rationality then **by making the rational rule accessible**, we should observe a **correction** of the framing effect.
 - Conversely, if the rational rule is **not used** as a guide by the subject, **while the rule is available**, this implies that the logical invariance principle is dominated.

«Norm retrieval» test (2 out of 3)

- In practice: PF and NF are presented simultaneously; then **the experimenter informs the subject of the logical equivalence** through a purely descriptive speech. [▶ more detail](#)
- Then the subjects are asked **in this order**:
 - Does the average person judge that the two problems are equivalent according to you?
 - Do you judge that the two problems are equivalent?
- Rationale for this dual design: **Bypass a possible «persistence bias» in the response.** [▶ More detail](#)

«Norm retrieval» test (3 out of 3)

- We are interested in the **reaction of a subject when realizing she was previously inconsistent.**
- A «*feeling of irrationality*» - «***epistemic discomfort***» - might exist.
- We envision to pin down such hypothetical feeling through an fMRI procedure.

Epistemic discomfort

- We are currently running an experiment resting on the «*bat and ball*» cognitive defect - massive rate of error while performing the cognitive task. One of our aims is to **capture the feedback process when the subjects realize their error**.
- Using these fMRI data we might **recognize** in the FE experiment **the same** specific neural substrate(s) activated after realizing a cognitive defect.
- Sort of «***epistemic discomfort***» when one realizes she was inconsistent.
- Look at the **contrast between the subjects who corrected** their previous inconsistent choices, **and the steady profiles** (steadily inconsistent and steadily consistent).

Two competing hypothesis to pin down the moral dimension

- If our competing hypothesis is true - sensitivity to the emotional-moral content of the task, which are **the mechanisms involved?**
- **Two alternative theories:**
 - A **purely emotional one** involving Dual Process Theory's automatic-experiential system - *S1* in Kahneman terms.
 - A **cognitive-based** one involving Dual Process Theory's analytical regions - *S2* in Kahneman terms.

The emotional hypothesis

- Choice of the gamble in the negative frame stems from a «*gut feeling*».
- Sort of emergency signal: «*I can't do that!*».
→ Specific activation of **limbic substrates** involving S1 should be observed when facing moral-content scenarios, vis-à-vis neutral ones.

Morality as broadened rationality

- The alternative hypothesis: the **trade-off** between the two options is **more complicated** within the negative frame as compared with the positive frame:
 - Within the positive frame, risk aversion dominates the choice: evaluation process boils down to choose a «*satisfying solution*», the «sure gain» (**bounded rationality**).
 - Conversely a **broadened rationality** is involved within the negative frame because the outcome is morally unacceptable to wit **trade-off between uncertainty and outcome**. [▶ Flesh out](#)
- Specific activation of cognitive substrates involving S2 should be observed.

Sketch of model to explain FE by introducing moral-emotional content (1 out of 2)

- «Broadened» rationality allows a difference of valence between the two equivalent – from a logical point of view – outcomes in FE.
- Use of a framework inspired of L L Lopes' theory of «security needs» and Prospect Theory's theory of status-quo.
- Define a «security seeker» as a subject who systematically outweighs the worst outcome in a gamble (sort of **max min behavior**) → The security seeker will choose the «*sure thing*» instead of the gamble in FE (to put it another way, she is risk averse).
- Hence if we just consider the «security needs» dimension (natural saliency of the worse outcome within a gamble), will predict risk-aversion behavior whatever the frame.

Sketch of model to explain FE by introducing moral-emotional content (2 out of 2)

- Now let's consider the emotional-moral dimension.
- Claim: the «*sure thing*» in the negative frame of asian-disease is a morally unacceptable outcome. The representation of myself's causing human death has a **higher valence** (namely stronger negative psychological consequences) than the same logical output in the positive frame.
- Optimal choice for a «security seeker»:
 - If the «*sure outcome*» is acceptable then the trade-off process is simple: choice of the sure thing. → **Bounded rationality**.
 - Conversely if unacceptable the cognitive process goes beyond. → **Broadened rationality involves a cognitive conflict**.

The retrieval test

- The experimenter explains the logical equivalence through a **purely descriptive speech**, namely she **does not** impose the rational norm:
«One might consider that the two problems are equivalent, to the extent that from a purely logical point of view, they are equivalent. Indeed, [here the analytical proof]...»

◀ back

These particular design is to reveal the «true» preference of the subjects

- We are aware of a **potential problem** associated with the «retrieval test»: if the subject was previously inconsistent in the FE experiment (namely provided a response in PF that is logically inconsistent with the one she gave in NF), might be inclined to provide **the same response pattern**; otherwise will suffer **psychological dissonance** → possible **Persistence bias**.
- We claim that **the first answer** (to the question: «*Would the average person judge that the two problems are equivalent?*») mirrors the «true» reaction of the subject (to wit the response without the dissonance effect); if **gap** between the first answer and the self-perspective → piece of evidence towards the existence of **psychological dissonance**.