

Framing Effects and Modal Illusions: How to Characterize Framing Effects as An Intensional Phenomenon?

Sacha Bourgeois-Gironde (sbgironde@hotmail.com)
ENS-lsh, Lyon¹

Empirically observed cognitive biases raise a challenge for the more formal and normative approach to rationality. Logicians, when facing these cases, can be led to reflect on some basic principles of theirs. Framing effects have been defined in terms of an intensional phenomenon (Schick, 1997)². Framing effects can basically be defined as such: they occur when equivalent descriptions of a decision problem give rise to divergent decisions. Framing effects come in various guises: some possible risky choice, some goal, or some attribute of an object or a situation is framed and yields a framing effect¹. Different interpretations have been proposed for framing effects as well as a drastic recasting of decision-theory. My more specific intention here is to try to understand in which sense framing effects constitute an instance of intensionality, i.e. a violation of a principle of extensionality, or invariance, in decision-theory and to what extent this form of intensionality is related to other phenomena of non compliance with extensionality in logic and cognition. Arrow (1982) seems to be the first to characterize framing effects as displaying intensionality, as D. Kahneman in his Nobel lecture reminds: “Framing effects violate a basic requirement of rationality which we called invariance (Kahneman and Tversky, 1984) and Arrow (1982) called extensionality”.

I am especially interested in the correlation between logical and informational lucidity and the possibility of framing effects. My point is to reflect on correlations between levels of subjective perception of logical and informational equivalence and decisions that obviously diverge from available knowledge. Traditionally it is to the extent that some piece of information or the relation between several co-reportive pieces of information is not cognitively processed that a diagnosis of intensionality is in order. When facing the extensional equivalence between P1: « 400 survivors out of 600 contaminated people » and P2: « 200 survivors out of 600 contaminated people » (speaking of the same group of 600 people), some psychological and pragmatic reasons may lead the subjects to focus on one piece of information (e.g. 200

survivors), in their choices or expressions of preferences, to the detriment of the other (e.g. irrespective of the complementary number of dead).

Frederic Schick suggests a strong notion of intensionality for framing effects as subjects being aware of the logical equivalence of two propositions may avoid processing them as co-extensional or as “co-reportive” in valuation and decision contexts. However it is hard to make a pronouncement about the degree of acknowledgment of co-extensionality by subjects in the absence of empirical data on that question. Under Schick’s definition, framing effects imply that even though the extensional equivalence of two propositions P1 and P2 is subjectively realized, they are not substitutable in choice or valuation contexts. A different approach would have it that even though logical equivalence is perceived between P1 and P2, P1 and P2 do not convey the same information as they bear distinct connotations. For this reason, framing effects do not exemplify any form of intensionality. This is the case if logic and information or perception of logical equivalence and perception of information equivalence can be wholly separated in the study of our phenomenon.

The defender of framing effects as significant violations of extensionality faces this dilemma: either he postulates a strong form of epistemic lucidity (perception of logical and/or informational equivalence) on the part of the subject – and it is hard to see how divergent decisions do not confine, then, to the irrationality of his choices and preferences –, or he, so to say, trivializes framing effects by saying that they systematically collapse to the justified perception of some informational discrepancies. There is, however, a way out which precisely stems back from the classical notion of intensionality. Information and logical behavior of sentences are related in the sense that bits and states of descriptive knowledge license some inferences and substitutions which others don’t. Implementation of logic depends on epistemic lucidity, but, given some accurate analysis of the intensional phenomenon at stake, trivialization doesn’t necessarily ensue even when subjects are lucid about logical equivalence of the statements which they consider because of the possibility of cognitively divergent aspects of the same piece of information conveyed by those statements in the subject’s mind. Sense or connotation, to adopt a classical philosophical terminology, can be cognitively divided as the study of so-called modal illusions has shown in the

¹ I thank Raphaël Giraud for ongoing conceptual discussions in view of the clarification of the topic of the intensional character of framing effects. I thank Andrea Armeni for his impulse to submit these ideas to the Stresa conference.

² F. Schick, *Making Choices : A Recasting of Decision Theory*, Cambridge University Press.

philosophy of mind. The more specific purpose of this paper is to envision some analogy between framing effects and modal illusions, in order to embrace the former with the resources of intensional semantics.

I will retain this broad characterization of framing effects: it is not the case that « P1 is logically equivalent to P2 » cognitively entails that « P1 is indifferent to P2 ». This is the phenomenon which I will try to theoretically account for. Labeling it a violation of the invariance principle of decision theory stresses a deep connection between extensionality in logic and indifference in choice and it consequently confers on classical logic an implicit normative scope in decision theory. Taking seriously into account the intensional phenomenon of framing effects invites to partially relax this normative constraint. It also invites to a retrospective look over the family of intensional phenomena. Intensionality has been traditionally linked to the cognitive division of information in a subject's mind. What is left of intensionality when complete information is in principle immediately available to the subject, as Schick suggests with his condition of epistemic lucidity.

1. Framing effects, ambiguity and epistemic lucidity

Framing effects are of various types. Typologies have been established by several authors. We can retain the one by Levin, Schneider and Gaeth who classify valence framing effects, i.e. variation of positive and negative elements in concurring descriptions, among three main categories. They propose correlations between what is affected by framing effects and what is framed. If what is framed is a risky option, what is affected is the willingness to take a risk. The framing of risky choices has been largely documented by T&K (1981), in particular through the paramount example of the ADP problem:

Imagine the U.K. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the program are as follows :

- A. If program A is adopted 400 people will live
- B. If program B is adopted, there is a 2/3 chance that 600 people will live and 1/3 chance that nobody will live
- C. If program C is adopted, 200 people will die.

D. If program D is adopted, there is 1/3 that nobody will die, and 2/3 probability that 600 people will die.

Other framings, according to Levin and al., concern the attributes of an object or the characteristics of an event or a situation and affect their evaluation. This type of framing is very useful to assess the hypothesis that framing effects depend on implicit reference points or baselines with which objects, events or situations are in fact implicitly contrasted. For instance if a medical treatment for an epidemics is presented as leading to a 65% survival rate or a 35% mortality rate, the first positive framing might well be accepted instead of the second because of some implicit reference to a lesser efficiency of a former similar treatment also framed in terms of a survival rate. Valence and implicit reference points jointly explain, in this case, the framing effect usually observed. A third type of framing is the framing of a goal which affects, according to Levin and his colleagues, the persuasiveness of communication. We will be mainly interested, in the rest, by attributes and risky choices framings.

A framing effect then combines ways of describing or presenting options, objects or prospects, which are normatively equivalent, and divergent levels of acceptance or commitment for some of those framed elements. We generally conclude from the higher acceptance of an option that it is the option preferred by the subject, the one to which she really commits and she would choose if actual choice were in order. We further traditionally conclude that given that those commitments are normatively inconsistent, the preferences of the agent do not exemplify rationality. However we can be reluctant to draw this inference from acceptance of a description to inconsistency in the set of preferences as long as we haven't justified how levels of epistemic acceptance collapse to an ordering of preferences. If one wants to take seriously the lessons of observable framing effects, of one type or another, one must be in a position to state that the agent's ordering of preferences is affected, which is, allegedly, a different diagnosis from saying that she has inconsistent or incomplete beliefs or epistemic attitudes; hence the importance to define and locate the intensionality of framing effects. If the information conveyed by the phrasing of the various options is completely processed by subjects, then it is reasonable to say that framing effects show an inconsistency or even a kind of irrationality in choice and preferences, which would require a proper explanation. If, otherwise, one can suspect that information is not homogeneous or processed as such by subjects, intensionality only affects choices within the context of the agent's epistemic attitudes and one has thus vindicated the rationality of preferences. The picture, though, might be a little more complex: it is not possible to clearly distinguish between processing of

information and expression of preferences, and this mere basic cognitive confusion may be at the source of framing effects. While this is the most basic explanatory hypothesis I entertain, I will have to take into account a few more pinpointed hypotheses about how information correlates with framing effects.

It is interesting to note that some researchers would not recommend ADP as a foremost and a good instance of a framing effect. Levels of informational explicitness between descriptions A-B-C-D are ambiguous. The lack of homogeneity could explain the divergent choices of subjects more than a lack of perception of informational equivalence when it is not clear that there is some. Mandel stresses, then, how the control of informational ambiguity between the several options might decrease the extent of observed framing effects in the case of the ADP³. A situation in which information is equally transparent or ambiguous for every proposed alternative might not completely reduce, though, framing effects because other factors can play a role. Among them, researchers generally emphasize salience and/or emotive value associated with the valence of descriptions. When risky choices are framed in terms of gains or losses, associated emotions may naturally explain why the behavior of subjects is not symmetrical. When some words like « dead » or « survivors » are used, salience of attributes, or associated mental pictures, can explain the discrepancies in choices and judgments too. Now our main question, here, is what to qualify as « informational » and what role information plays in the characterization of framing effects as an intensional phenomenon. It is only in connection with information, and not with emotion, that talk of intensionality of framing effects makes sense.

Some disambiguation, of terminology this time, must be in order. The usual characterization of framing effects is put in terms of *equivalent* descriptions of a decision problem yielding systematic divergent decisions. What kind of equivalence? There are at least two possible candidates: logic and information. What does it mean for descriptions or propositions, in a decision problem, to be either logically or informatively equivalent? How does either kind of equivalence or choice bear on the rationality of decision? One is led here to contrast, in the traditional Fregean way, reference, or extension, and sense, or intension. Reference is what contributes to the extensional characterization of the full sentence, the fact that it reports, or here, the fact that would be the case if an act in view of its establishment were taken. Sense is what

³ Mandel, D. R. (2001). Gain-loss framing and choice: Separating outcome formulations from descriptor formulations. *Organizational Behavior and Human Decision Processes*, 85, 56-76.

descriptions used to fix extension independently mentally convey. The striking effect is that what descriptions convey may lead the agent to think he prefers to establish one fact rather than the other while there is only one. But descriptions also help to fix the extension of the terms contained in sentences. So one might prefer to contrast logical equivalence and informational equivalence of two descriptions or sentences as, for instance, Craig McKenzie does: « A pair of statements is logically equivalent if each member of the pair necessarily entails the other. (...). When there is no choice-relevant background condition C about whose probability a listener can draw inferences from the speaker's choice between frames A and B, we say that A and B are « information equivalent » »⁴. In that definition the two notions of equivalence are better distinguished than in the Fregean approach: description does not help to fix extension, it is kept independent from the logical component of the problem, but it can continue to provide the required explanation for divergent choices.

There is a clear dividing line between logic and information in McKenzie's account. Logically equivalent options may lead to divergent choices insofar as leakage of information from the different descriptions is itself different. From the fact that a speaker chooses the frame A, C may be inferred, which is not the case from a frame B. That a less successful treatment was previously administered in cases of some epidemics may be inferred from the description of a medicine in terms of a 65% survival rate; it is maybe less obvious to draw the same inference from the description of the same medicine in terms of a 35% mortality rate. McKenzie's distinction relies upon a pragmatic charity principle: in case of apparent violations of invariance, some relevant information C must lie in the background: « Suppose (...) that speakers, choosing between uttering 'A' and uttering 'B', are more likely to utter 'A' when some background condition C holds than when C fails. In that case, a listener who hears a speaker say 'A' can safely infer a higher probability of C being true than if the speaker had said 'B', that is, $p(C|\text{speaker says 'A'}) > p(C|\text{speaker says 'B'})$. If knowledge about the background condition C is relevant to the choice at hand, then the speaker's utterance of the two logically equivalent statements A and B may with impunity lead to different decisions ».

The difference between Frege's and McKenzie's approaches to intensionality can be phrased in the following way. For Frege different senses might determine the same extension (i.e. the same fact or object corresponding respectively to the different sentences or descriptions). But a subject being unaware of this equivalence prevents him from recognizing the logical

⁴ Personal communication, under press.

equivalence between sentences and, thus, from drawing the same inferences from each of them. The phenomenon is then as such: lack of perception of co-referentiality (or generally co-extensionality) leads to the non-application of a principle of logic (the substitution of co-extensional expressions in sentences that contain them). What is interesting with framing effects, as McKenzie understands them, is that we have somewhat the reverse cognitive situation: even though logical equivalence is perceived and admitted by the subjects, one description may not be substituted to the other in choice contexts (or even in contexts in which subjects are led to merely express their 'epistemic' preferences), one inference, to the presence of some piece of background information, drawn from one sentence may not be drawn from another. This is due to connotations or informational implications conveyed by such and such descriptions (even when they are extensionally equivalent). So instead of a lack of cognitive connection between co-reportive pieces of information yielding a violation of logic - in the classical post-Fregean notion of intensionality - we have, in McKenzie's explanation of framing effects, the perception of logical equivalence not being a sufficient cognitive ground to operate normative substitutions of coextensional descriptions.

2. Intensionality

It just then imparts some thoughts about how this notion of intensionality, with which philosophers are familiar, really applies in the discussion of framing effects. In fact its first interest is to help to raise some fundamental points about this distinction between logical equivalence and information equivalence. Logical equivalence is an "objective" datum while information equivalence is a subjective one - or this view is prevalent and makes most people prone to locate the source of normativity in decision theory in logical equivalence. Let's introduce a notation for these two notions of equivalence: xIy for x is informatively equivalent to y and xEy for x is logically equivalent to y . The prevalent hypothesis on the source of normativity becomes: $xEy \Rightarrow xIy$. Now one can distinguish three grades of intensionality. I use a non specific epistemic or cognitive operator Bp to note p is epistemically accepted by the agent, with no more ado, for the time being, about what 'epistemic' exactly means in this context:

- Fregean intensionality: $\text{not } B'xIy' \Rightarrow \text{not } B'xEy$
- McKenzie's intensionality : $\text{not } (B'xEy' \Rightarrow xIy) = \text{FE}$ (framing effect)
- "hyperintensionality" : $\text{not } [(B'xEy' \text{ and } B'xEy') \Rightarrow \text{not } \text{FE}]$

Fregean intensionality is captured in terms of the lack of epistemic acceptance of two pieces of information as being co-reportive preventing the perception of logical equivalence. McKenzie's notion of intensionality amounts to the non systematic entailment of informational equivalence from the acceptance of logical equivalence between options. This is McKenzie's basic account for framing effects.

The third grade of intensionality suggested is quite speculative and confines to the irrationality of subjects who would display the corresponding behaviorⁱⁱ. It amounts to the epistemic perception of logical and informational equivalence not preventing framing effects to occur. If such a phenomenon occurred, it would show a more profound discrepancy between processing of information and logic on one side and choice on the other side than the two former notions of intensionality. So we could be content to say that the contrast between logical and informational equivalences captures a good deal of what could be meant when applying this notion of intensionality to framing effects. However, we cannot completely discard the third grade of intensionality as it reveals some inherent risk in McKenzie's position. The problem is that no phenomenon of framing effects can be connected with irrationality under McKenzie's view if there is always a reference-point or a background information C that a listener can infer from the choice of some frame by a speaker. However we cannot be sure that framing effects are not a symptom of a more irreducible form of epistemic inconsistency or even of a form of meta-inconsistency between epistemic acceptance on one side and choice or decision making on the other side. This is the hypothesis we clearly follow now when we try to establish an analogy between framing effects and modal illusions.

3. Framing effects and modal illusions

One way of stating in which sense framing effects instantiate intensionality, that is without the presupposition that explicit or latent equivalent pieces of information are not perceived as such, is, to borrow from the philosophical literature⁵ and say that the subject is victim of a 'modal illusion'. Let's spell out that suggestion and see how the agent's rationality can be preserved when choice situations are analyzed from this modal angle, even though his preferences *seem* to conflate with his epistemic states.

Modal illusions have been recognized in the philosophical literature as cases in which subjects make systematic

⁵ See the anthology by Gendler, T. S. and Hawthorne, J., "Conceivability and Possibility", Clarendon Press, Oxford 2002, especially the introduction where the philosophical notion of modal illusion is presented.

mistakes on the modal status of some classes of propositions. Some proposition P is necessarily false while it will be usually deemed as possible. Water is necessarily H₂O while some will judge it could have turned out that water is chemically different from H₂O. Metaphysical and semantic reasons have been given to justify the claim that some proposition X displays a necessity and that a statement to the contrary constitutes the basis of a modal illusion. But not much has been said on the psychological phenomenon that underlies so-called modal illusions. While it is not our purpose to get into this analysis here, it is interesting to focus on this more psychological aspect. It is not clear that subjects are absolutely blind to the modal status of the original proposition they consider. An alternative explanation is that while they recognize the actual necessity of, say, water being H₂O, they project themselves in some counter-actual situation in which the facts are less metaphysically determined. It is a case of epistemic indeterminacy that does not necessarily points to any kind of ignorance of what is actually the case. Philosophers have used formal devices to give a simple account of this epistemic shift. They used what is named “2-Dimensional Modal Semantics” (henceforth 2DMS). 2DMS predicts that the actual modal value of a statement can be ‘necessary’ when epistemic focus is borne exclusively on the actual world, while it can vary when epistemic focus is made less determinate and can encompass other reference worlds than the actual world. Let’s restate this modal variation and compare it with what we said so far on framing effects, information and epistemic lucidity.

Let us start by looking at an isomorphism between framing effects and modal illusions.

Framing effects share this common structure:

A is equivalent to B
(Two options are normatively equivalent)

And still:

A > B
(A is behaviorally preferred to B).

Modal illusions exhibit the reverse isomorphic structure:

A ≈ B
(A seems epistemically equivalent to B)

While in fact (from an accepted normative standpoint underlying modal judgments)

A >* B
(A is a closer possibility than B)

Let us return to the previous example of a modal illusion where a subject experiences no epistemic qualms in accepting the possibility that water is not H₂O while the subject also accepts in principle some normative standpoint that would bring forth the a posteriori necessity that water is H₂O.

A first point to notice is the possibility that, given this isomorphism, the two situations of modal illusions and framing effects can be dealt with, to some extent, through a common framework, 2DMS. 2DMS is a logical framework that allows for a modalized sentence to be interpreted relatively to pairs of worlds rather than to worlds *simpliciter*. Any sentence, then, can be interpreted relatively to the actual world and relatively to any other world taken as actual; its counterfactual content is then contrasted with either of the worlds considered as actual, whether actually actual or not. So in jointly admitting the stance that “Water could be different from H₂O” and that “Necessarily water is H₂O” the 2DMS reading of the apparent inconsistency holds that in the first case the subject envisions as its reference-world in one case the counter-actual world, and in the second case the actual world.

2DMS disambiguates reference-worlds for the interpretation of modal sentences. A counter-actual world is associated with some private epistemic state of the subject: it is the world to which the most idiosyncratic beliefs of the agent can be made relative. Even if this private state can be found in more than one and even a majority of subjects, counter-actual worlds reflect subjective meanings people associate with the acceptance of some sentence as possible while it is normatively impossible. To insist on this subjective side of modal beliefs: people entertain an impossible state of affairs and they find private meanings that describe a possible scenario they circumstantially take as their reference-world. Now, intensionality would be limited, in the case of modal illusions, if the worlds people select as their points of evaluation could be systematically related to the actual world as we know it. This, however, can’t be the case. In order for 2DMS to properly apply, no backward reference to actuality (actual actuality) can be made in the purview of a counter-actual standpoint, lest the latter immediately cease to be a full-fledged counter-actual standpoint. A resort to a 2DMS explanation of modal illusions does not per se allow one to have it both ways: private and yet explicitly informational. It can still be taken as a good device of disambiguation between private and public meanings when the two are, unconsciously or consciously, kept epistemically apart by the subjects. This ideal epistemic separation is what makes the 2DMS treatment of modal illusions potentially congenial to the study of the kind of intensionality framing effects are endowed with.

In both framing effects and modal illusions, subjects make themselves locally blind to some informational equivalence between two propositions and display epistemic preference for one over the other. One can connect framing effects to modal illusions in the following way: favoring one possibility can be interpreted as taking this possibility as implying a reference-world different from the one the other possibility would imply and failing to see that the two possibilities in principle refer back to the same world. We have here a kind of in-built modal illusion in the following sense. It is as if, indeed, the options expressed by two extensionally equivalent propositions, in a framing effect, were not counterfactually related one to another by the subject. It is as if their equivalence was not perceived and as if, for the subject, the two options reflected two essentially different things rather than two distinct descriptions of the same thing. In other terms the modal illusion is here deepened into a form of objective illusion. The various options presented in a choice situation in which a framing effect arises are perceived as different objects even more than as different descriptions of the same object. This deepened modal illusion seems to provide a tentative diagnosis of what kind of errors could be involved in framing effects.

One can express the reverse isomorphism between modal illusions and framing effects in more than one way. In a modal illusion we need a disambiguating framework in order to realize that we use the same words to refer to different objects. In a framing effect we seem to disambiguate when we need not because the terms (the descriptions associated with the options) actually refer to the same object (option). When an agent rejects the 2DMS diagnosis for modal illusions because one

maintains that he uses the terms to refer back to the same thing (for instance to water *as we know it*). One makes what could be called a hyperrigid use of those terms: he wishes to maintain reference of those terms in context where such reference is not available. When victims of framing effects, on the contrary, agents do not see that descriptions bear on the same object, that is, they fail to see co-referentiality and show an extreme form of intensionality (in not allowing for the substitution of equivalent descriptions).

This isomorphism between modal illusions and framing effects may be the symptom that those two types of cognitive biases may represent two faces of a common phenomenon: the role modal cognition plays in decisions as well as conceivability. This cognitive ability has some particular features and when conflated with other cognitive abilities such as reasoning or processing information, some apparent mistakes or failures may appear. One needs not diagnose deep irrationality though, to the extent that we are able to identify the biases and to disambiguate modal reference points (rather than informational reference points), we rather see the typical contribution that modal cognition make to decision processes, a contribution that was shadowed by more purely informational approach to decision processes and biases attached to them.

ⁱ See the typology by Levin, Schneider and Gaeth, ***

ⁱⁱ Tests in order to see how subjects behave when faced with pairs A/C and B/D (within subjects tests) should be proceeded.