Judgment and Decision Making, Vol. 3, No. 4, April 2008, pp. 355–363

A note on neglect defaulting

Howard Margolis*
Harris School of Public Policy
University of Chicago

Abstract

I introduce the notion of "neglect defaulting," which labels the propensity to neglect possibilities which are ordinarily sensibly neglected. In familiar contexts we are well-tuned to recognize when to override the default. But outside the range of familiar experience — here in the artificial context of puzzles — these ordinarily benign defaults can make it difficult for even sophisticated subjects, such as readers of this note, to avoid responses which on reflection will be seen as obviously mistaken. A detail of particular importance is that, although subjects are easily prompted to take one step in the direction of reaching a sound response, the tendency to then neglect to consider that another step may be needed is remarkably strong. In each of the five examples the needed but usually neglected second step is quite trivial. Concluding remarks point to consequences for larger questions outside the range of familiar experience, in politics and other contexts out of scale with everyday experience.

Keywords: puzzle problems, Bertrand's box, Wason selection task, economic games, heuristics, defaults.

1 Introduction

The *neglect defaulting* I discuss here can be viewed as a generalization of the familiar phenomenon of base-rate neglect. The notion overlaps the anomalies of Kahneman and Tversky's "heuristics and biases" program. But it extends to anomalies rarely or never discussed in that context, and in ways that can sometimes clarify what is happening in anomalies within that program. As with base-rate neglect, the neglect is not universal (subjects sometimes do not neglect base rates), nor is it always an error (sometimes neglect of base rates is normative). But in contrast to positive heuristics, neglect defaults are not comfortably seen as *de facto* rules-of-thumb usually good enough to yield a sensible response. "Neglect base rates" does not seem a sensible rule of thumb, though very often it is in fact sensible (Einhorn & Hogarth, 1986). It is one of a considerable class of cognitive effects which might be called "negative heuristics", but the more explicit label of "neglect defaults" is less likely to be misunderstood.

Neglect defaulting should be seen in terms of a very general Darwinian point. Creatures who survive must come equipped with defaults for what to do when they would otherwise be unsure what to do, guiding responses when familiar cues are sparse, or weak, or hard to pick out, or conflicting. A conspicuous example would be the fight-or-flee situation. This is pervasive in nature. Indefinite hesitation is not a viable option. And even an elephant would probably do best if equipped with a *flee* default when in doubt, though presumably an elephant would not often be in doubt.

What makes sense of cognitive shortcuts in general is that they conserve attention when the delay otherwise required would be too costly to tolerate (as with fight-or-flee situations), or simply useless, or unlikely to do much better than a quick rule-of-thumb response to the circumstances at hand. But for *neglect defaulting* the economy of use on any particular occasion is usually negligible. Rather, what accounts for neglect defaulting is that the occasions for the default responses are so very common. The defaults always concern classes of choices where, without neglecting by default almost all occasions to expend attention, a person would be overwhelmed by hesitations, even if only brief hesitations. These are hesitate/*proceed* defaults. As long as the default is in place, we rely on our "blink" intuition and *proceed*, not hesitate.

We constantly make choices, mostly trivial, mostly in fact not reaching the level of conscious attention. We also constantly encounter opportunities to hesitate and reconsider whatever intuition is at hand. So if we had to adopt a default rule about whether to stop and think longer about a choice when we have no sharp indication one way or the other, we would be crippled unless that rule made us unlikely to hesitate. If I am walking across a minefield, it will not take much to get my attention, but in an ordinary field I will not stop to consider how to maneuver around every bush I encounter. I know it is possible that if I looked more carefully I would see a rusty nail or maybe

*Address: 1155 E. 60th Street, Chicago, IL 60637, Suite 164. E-mail: hmarg@uchicago.edu

even a rattlesnake. But I move on, disregarding such possibilities unless some cue is salient enough to displace the *proceed* default. Random choices will be almost always be good enough, while spending time mulling every choice would be a disaster. But inevitably cases will arise in which proceeding rather than hesitating will turn out to be a mistake.

In Margolis (2007, Ch. 6) I describe and illustrate various forms of neglect defaulting (modus tollens neglect, ordinary language effects, and several others). Each is some functional equivalent of a hesitate/*proceed* neural switch whose default position is "proceed." The effect of the default is then to neglect some possible occasion for stopping to think about how an incremental piece of information might be relevant to a choice at hand — in the context of a puzzle, how to answer the puzzle correctly. Each neglect default allows a person to proceed rather than hesitate unless a sufficient jolt pushes the default switch to hesitate. But sometimes the context at hand is one where that neglect is adverse. It allows a person to move ahead with a choice which on reflection does not seem sensible, in a context where that could have been avoided if the default had not cut off a bit more thought.

But no one *chooses* to neglect an opportunity to reconsider a default. *The essence of the matter is that the default shuts out noticing any such possibility.* But, since someone merely contemplating the possibility of a stubborn and invisible default cannot see this, it is tempting to explain away anomalous responses by supposing that the problem was somehow more difficult than it might look. For suppose that indeed I have an opportunity to improve a choice. There might nevertheless be enough difficulty or complexity to the problem to make it hard or unprofitable to actually use that opportunity. If that is all that is happening, then neglect defaulting might have no serious consequences. So it might be tempting (as a way to rationalize neglect) to conjecture difficulties. And that is easy to do, since even inferences that we make routinely and easily will sound complicated and difficult if we try to spell out every detail of how to formally reach the inference. I can assure you that it is no trouble at all for me sneeze. But if someone could write out all the nerve signals and muscle movement required for that, you might wonder how I can possible manage it all. The examples coming next are organized to pre-empt that sort of denial.

In terms of neglect defaulting, base-rate neglect is a special case of what I've called "twoness" effects (Margolis 1996, ch. 4), which is a general tendency to neglect a need to balance between polarized perspectives, here local vs. global cues.[1] Base-rate neglect is the particular possibility that assessments of recent or local experience might be improved by adjusting in light of some item of long-term or global experience. In familiar situations a person is likely to be very good at noticing when that default should be overridden. But, in unfamiliar contexts, such as in the artificial context of lab puzzles, or even in a familiar context seen in an unfamiliar way, arranging a sufficient jolt to offset the default may not be easy. Base-rate neglect, since it lends itself to formal discussion and tightly designed experiments, has taken on a considerable life of its own (Koehler, 1996), but in terms of neglect defaulting it is better seen as an exemplar of a broader tendency. So if neglect defaults might be much harder to dislodge than we would consciously see as sensible, then base rates would sometimes be ignored by default, even when it is strikingly clear that they shouldn't be.

2 Five examples

What follows are five simple choice problems of very different sorts where what seems to be adverse neglect defaulting can account for puzzling experimental data.

2.1 Bertrand's box problem[2]

Three cards are in a box. One is white on each side, and another is red on each side. The third is white on one side and red on the other. So there is one white/white card, one red/red card, and one white/red card. Without looking, you take out one card, and lay it on the table. Consider these arguments:

(A) Of three equally-likely cards, two are same color cards, and only one is mixed color. So before you see the color, the chance is 2/3 that the card you picked is a same color card.

(B) Of three equally-likely red sides, two are on the red/red card, one is on the white/red card and none are on the white/white card. So if the color you see is red, the chance is 2/3 that the color on the other side is also red.

(C) If you agree with one of these arguments but disagree with the other are you contradicting yourself?

The marked tendency is to find A easy to accept, but B hard to accept. And among the large fraction who report confident intuitions that A is right but B is wrong, no one I have encountered in many trials has ever noticed, even when prompted by C, that in fact their intuitions are contradictory. So you might be caught also! But a reader caught by this intuition trap can easily resolve the matter

[1] Other pairs of polarized alternatives, where almost always some balance between the alternatives would improve judgment would be (from statistics) errors-of-the-first-kind vs. errors-of-the-second-kind, (from social psychology) attributional vs. situational interpretations of behavior, (from policy assessments) significant vs. negligible risks, and many more. See, e.g., Margolis 1996, pp. 81–6.

[2] This is a variant of a 19th Century puzzle which became prominent in JDM-related papers through Falk & Bar-Hillel (1982), with many subsequent discussions.

by a simple experiment which makes it easy to see that these intuitions, in combination, are in fact nonsensical.[3]

The most common intuition going against the B claim is that the chance the card you picked is red on the bottom when red on top is 1/2, not 2/3. But everyone who has that intuition also sees the chance the color on the bottom is white if the top is white as again 1/2. If so, then even before you see the color, you can say that the chance is 1/2 that the color on the bottom of the card will be the same as the color on top. But from A you also believe that before you see the color, the chance is 2/3. The contradiction is logically trivial, but somehow hard to see even facing the argument you have just been given. But anyone who puts himself in the physical situation is quickly cured of the illusion. There is never any need to run a large number of trials to correct the illusory 1/2 intuition. Once in the physical situation, after no more than a dozen trials it becomes obvious that argument B is correct. But unless you do the experiment, the illusion may be remarkably stubborn.

Here is an account of what generates the illusion which will also apply to each of the four further examples to come. The characteristic form of an effective puzzle is that a chooser is introduced to a context (here the 3-card setup). Chooser is then confronted with a new piece of information: here, "the side is red." Subject responds, given the basic context plus the prominently-presented new information.

And suppose neglect defaulting occurs like this: Given the salience of the new information some immediate intuition will be prompted. But after this minimal response, suppose the neglect default easily slips back into place. Strikingly effective puzzles would then be just those formulated in a way that this "one-step" response is unsound. We would have a strong selection effect. The puzzles that draw the attention both of people who like puzzles and of scholars who like to study puzzling choices would be just those which present some critical cue which requires more than one immediate intuition to yield a correct choice, but the cue is not strong enough to yield more than the momentary, one-step, escape from the neglect default.

We might think of the default as operating like a toggle having an intermediate position. Any mild jolt is likely to be sufficient to push the default momentarily aside, prompting some one-step intuition. But then there is reversion to the default. A stronger jolt is needed to get past the one-step position to full escape from the default, allowing reflective thinking. Or, likely to be much easier, perhaps the issue can be presented in a way that mitigates the one-step propensity by in some way putting the usually-missed next step in front of the hard-to-miss first step.[4] Examples to follow will try to make clear just how that works.

If I say "canary" you will have an immediate sense that it sings or that it is a bird. But we do not have such an immediate sense that it lays eggs, though if asked it does not take long to agree. If the context was a puzzle where "egg" was the correct response, the chance of getting it right might be enhanced if the framing made "bird" the more likely one-step response, and the problem might become very difficult if the likely one-step response was "sing." Simple relations of this sort let us contrive variants of puzzles that yield sharply different responses to framings that (until we've seen the effect) apparently differ hardly at all. I will give an example for each of five puzzles discussed, starting with the 3-cards puzzle already at hand.

The "side red" cue might prompt a subject to note that (a) the white/white card has no red side. Or it might prompt (b) two of the three red sides are on the red/red card. Someone who is prompted to (a), but the response is one-step, does not notice (b), though as logical tasks both are about equally trivial, and indeed in question B is explicitly in sight. But if that happens, chooser is left with the prior intuition that the available cards are equally-likely, except that there are now only two, which generates the modal but mistaken intuition that the chance the color on the bottom is also red is 1/2. And anyone who has tried this puzzle on colleagues or students will be aware how stubborn this illusory response is likely to be. That 1/2 is the modal response, though with the also illusory 1/3 much more common than the correct 2/3, reveals that (a) is more readily prompted than (b). But suppose the puzzle was framed in a way that changed that. Then one-step defaulting would make it likely that (a) would be missed. So on this account we might look for some alternative framing that might make (b) as likely (a) to be the one-step response, or make it more likely that both would be noticed.

Fox & Levav (2004), for reasons unconnected to the discussion here, ran the 3-cards problem with Duke students, yielding an even lower than usual fraction of correct (2/3) responses. Barely more than 1 in 40 students at this strong university got this logically simple problem right. But Fox & Levav, exploring parsing effects, then focused attention on the *sides* of the cards rather than the cards themselves by emphatically making the question about the chance if a red side was seen it was on the card with two red sides.[5] This yielded a large improvement

[3] Mark three slips to make the RR, WW, WR set. Shuffle them and pick without looking. Keep track of the result over a series of trials.

[4] There is a kinship here with the "level k" analysis of strategic choice in recent behavioral game theory (Crawford & Nagore, 2007, Camerer et al., 2004). In that framing, k = 0 would be complete neglect, k = 1 would reach the one-step momentary escape, and k = 2 on would reach successively deeper levels of reflection.

[5] The cards were described as red1/red2, white1/white2, and white3/red3. The question asked for the chance that a red side show-

Table 1: A sample of the Charness/Rabin games.

Game	(Trials)	Choices:					A: Left/Right	B: Left/Right
CR9	(36)	A (450,0	*or*	B (350,450	*or*	450,350))	.69/.31	.94/.06
CR28	(32)	A (100,1000	*or*	B (75,125	*or*	125,125))	.50/.50	.34 .66
CR14	(22)	A (800,0	*or*	B (0,800	*or*	400,400))	.68/.32	.45/.55
CR18	(32)	A (0,800	*or*	B (0,800	*or*	400,400))	.00/1.00	.44/.56
CR26	(32)			B (0,800	*or*	400,400)		.78/.22

over 2.6% correct in their trials using the basic puzzle. But 2/3 was still not the modal response, which switched from 1/2 to 1/3, which is even further from the correct 2/3. So the manipulation was often effective in moving subjects away from the illusory 1/2 intuition to the correct 2/3, but it was just as likely to move subjects in the wrong direction, from 1/2 to even less correct 1/3.

One-step defaulting provides a simple explanation. With this strong manipulation, many more subjects (27%) now saw both (a) *and* (b). But by a wide margin most still saw (a) or (b) but not both. Despite the strong manipulation, 24% still chose "1/2", which would come from those who saw (a) but not (b). A larger number (35%) responded "1/3", which would come from those who saw (b) but not (a). For a chooser who got (b) but missed (a) would see the chance the red/red card was chosen as 2/6 instead of 2/3.

And a closely parallel account resolves an entirely different sort of puzzle next, and in a way that makes the one-step effect especially easy to see.

2.2 Charness/Rabin

Daniel Kahneman's Nobel prize was shared with Vernon Smith, who led the development of the novel field of "experimental economics." Most of these experiments from economists deal with markets or auctions, but many have dealt in one way or another with issues of social cooperation. Since norms of reciprocity are a feature of every culture, and in obvious ways promote social efficiency, many experiments in one way or another test for conditions where reciprocity norms succeed or fail. Table 1 shows a sampling from the 32 "simple games", mostly offering opportunities for reciprocity, reported by Charness & Rabin (2002) in a widely-cited 2002 article.

Reading across the entries, we have a game ID, the number of trials, then the choice offered the A player. The A-choice is between moving *left,* which yields the first set of payoffs, *or* pass the choice to B, who chooses between the pairs of payoffs on the *right*. In CR9, for example, if the A-choice was *left*, A gets 450, B gets 0. But if the A-choice was *right*, then the B-choice determines whether the payoff will be 350 for A, 450 for B, or the reverse. In all but CR26 — where there is no A choice — each subject plays the game twice, once in the A-role matched with an anonymous partner and another time in the B-role, matched with a different anonymous partner.

The columns on the right show the results, which are curious not only for the sample of games shown here but throughout the 32 games in the set.

The players see a tree diagram showing the game they are in, like the one for CR9 on the left of Figure 1. The A-choice is whether to move left or right in the top branch of the tree, without knowing what B will choose. The B-choice is whether to move left or right on the lower branch, without knowing what A has done. But B does know (since he is looking at the game tree) that the B-choice becomes relevant only if the A-choice turns out to be *right*. So the B-choice is about what to do if A has passed the choice. The A-choice is whether to make the B-choice relevant, or just take the payoff pair on the far left.

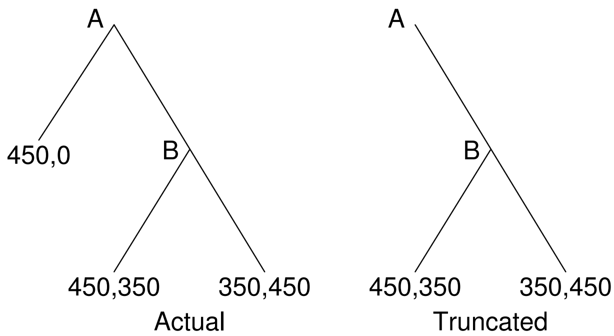


Figure 1: Charness/Rabin tree for game 9.

In CR9 a self-interested choice for A is easy. A can gain nothing from making the risky choice of *right*, since at best that would be no better than the 450 available for sure from choosing *left*. But a player who would like to help B has a risky choice. He can let his partner get 350 instead of 0, but if he passes the choice (chooses *right*) partner could betray him and take the 450 for himself.

ing after the pick was red1 or red2.

https://doi.org/10.1017/S1930297500000929 Published online by Cambridge University Press

In terms of reciprocity norms in every culture that would be really bad behavior, returning a generous act with a selfish one, but it could happen.

But almost one third do take that risk as A. They choose *right*, bringing the B-choices into play. And the B-choices (by players, recall, who are also making A-choices) turn out to be just terrible for A. In the CR trials, in the B-role 94% betrayed their partner! The fraction is so overwhelmingly large that it necessarily includes the great majority even of those whose A choice reveals they are people who not only prefer to be generous but are willing to run a risk to be generous.

The choices appear to reveal entirely different motivation for A and B in CR28, but the results are equally odd. Half the A choices now are nastily selfish, and reckless as well. As A, players are as likely as not to throw away 900 of 1000 B tokens for what might seem the remote prospect that in return B will reward them by giving them a gain (increasing their payoff from 100 to 125) rather than a punishment (cutting their payoff from 100 to 75). But that selfish and reckless move turns out to be profitable, since 2/3 of these same players (from the same subject pool so prone to betrayal in CR9) choose as B to be generous in responding to an aggressively selfish A choice.

But the game tree reveals a one-step character in these puzzling responses. Players predominantly make B choices that make no sense as contingent responses to A choices which bring the B-choice into play. Where A-choices are generous even though it is risky to be generous (in CR9), B-choices overwhelmingly betray generosity from A. But A-choices that are selfish (as in CR28) get mostly generous B-responses. Subjects grossly violate everyday experience and violate the norms that parents in every culture take care to inculcate in their children. Their choices conflict with numerous experimental results from both social psychology and economics in which marked reciprocity effects are routinely observed (as in Charness, 2004).

But if players are vulnerable to one-step defaulting, the game tree imposes an ordering which determines what B-choosers will neglect. B can notice the payoff he would have gotten if A did not choose right only after noticing that the choice he faces was presented to him by A. The *truncated* game tree on the right of Figure 1 is the only one-step sense of the situation that could occur. And truncation turns out to makes sense of the bizarre responses seen throughout the CR games and illustrated here with a few examples.[6] Again and again the data report "normal" responses to an illusory perception of the context, since all that B seems to responds to is the immediate awareness that A presented him with this choice. If that were the actual situation, why would a person choose "me 350, you 450" over "me 450, you 350"? Or why would a reasonable person stingily prefer "me 100, you 75" over "me 100, you 125." This is not, after all, a competitive game. The predominant choices are "normal" responses, but only to the truncated game (on the right in Figure 1), not to the game the players are actually in (on the left).

This reading is reinforced when CR14, CR18, and CR26 are considered. In each the B-choice is identical. But B is generous twice as often in CR14 and CR18, where there is an A-choice, compared to CR26 where there is no A-choice. Unmistakably B-choosers do not miss the bare point that their opportunity to make a choice depended on A presenting that opportunity to them. But also unmistakably they respond only to the truncated game-tree, neglecting entirely the alternative A could have chosen. B-choices are not in the least more generous to A in CR14, where A is taking a large risk to offer a generous choice to benefit B as against CR18, where A's choice is completely self-interested. Almost half the time as B, choosers betray a very generous A choice in CR14. More than half the time, they are generous to A in CR18, even though A has done nothing at all to elicit a generous response. Transparently, B-choices are neglecting what the B payoffs would have been if A chose left. B's are considerably less than half as likely to share their payoff with A in CR26 than they are in CR18, though A has done nothing at all to deserve that. In CR18 A reaps a splendid reward from merely passing the opportunity to give him half the payoff.

A curious aspect of the CR data is that the bizarre effects, going aggressively against both common experience and numerous experiments revealing marked reciprocity effects, have been completed ignored in the many citations of these games. Charness & Rabin themselves only note that reciprocity seems to be absent in their games, without further comment on the point. But the situation is in fact much stranger. It is not that on average subjects show null reciprocity effects. Rather there are well marked reciprocity effects, but they almost always (as in CR9 and CR26) go in the wrong direction. The null net effect is due to clear positive and negative reciprocity effects, each in the wrong direction for the context, but cancelling out overall. We might expect that such strong and unexpected effects would get a good deal of discussion. But what seems inexplicable if noticed easily comes to be ignored.

And a parallel to that arises next in connection with a puzzle (Wason's "selection task") so familiar as to risk prompting yawns as I bring it up. But Wason's puzzle yields a striking novel result when looked at in the light of neglect defaulting.

[6] There are also striking but subtler puzzles in the A-choices, which I cannot deal with here. But A-choices as well B are treated in detail in Margolis (2007, ch. 9).

2.3 The Wason selection task

As with the "base rate" issue mentioned earlier, "modus tollens" is academic talk for something routinely encountered without any such label. When one thing entails another, if that second thing is absent the first thing ought to be also absent. You need no formal study of logic to know that "if it's raining it is cloudy" lets a person suppose (modus tollens) that if there are no clouds it is not about to rain. But opportunities to notice totally useless modus tollens inferences are ubiquitous. Overwhelmingly, we just ignore those opportunities. Using an example much discussed among philosophers (Hempel's paradox), suppose I assure you that "all ravens are black." And I am wearing a green tie. Noticing something green, and observing that it is my tie, should you not gain an extra mite of confidence that indeed all ravens are black? For no one doubts that "all ravens are black" implies (you will follow even if you have never studied formal logic) that "non-black things are non-ravens." And my green tie indeed is not a raven, just as predicted by a theory that all ravens are black. So here is a bit of evidence supporting that theory. Philosophers debate how to handle this silly but logically impeccable inference.

But unless provoked by an academic discussion we do not notice this silly inference at all. Rather we are protected from wasting time noticing it only to have to waste more time concluding it was not worth noticing. It takes a ravenish non-black bird to provide us with an occasion where we might check for a case where a raven is not black. But it only takes an object which is not black to provide an occasion when we might check for a case where (modus tollens) any non-black thing is a raven. We could spend the day gathering such evidence at a prodigious rate and be no more confident than we were at the start that indeed all ravens are black.

Nor is the burden of inferring what might be implied modus tollens limited to the very occasional cases of sweeping generalizations like "all ravens are black." Mom says, "Dinner will be ready by the time you get hungry." So the kid expects that if dinner is ready, he will be getting hungry. But he is not burdened every few seconds over the next several hours with thoughts of "I'm not getting hungry, so dinner isn't ready."

So one of the likely hesitate/*proceed* defaults that enters this discussion of neglect defaulting is modus tollens neglect.

As with base rates, it is not that people untrained in formal logic are incapable of noticing and using everyday versions of modus tollens. Rather, on this neglect defaulting account, a tacit default blocks such inferences unless there are sufficient cognitively effective cues in the context to displace that default. The usual account of the Wason task turns in some way on how costly it is purported to be for a normal human being to use modus tollens, as illustrated by a comment like this from a scholarly volume of papers on the psychology of reasoning: "After several years of teaching this topic the current author believes he has finally learned the Modus Tollens rule" (Roberts, 2000). But if the failure to see the modus tollens inference is due to neglect defaulting, the difficulty is not at all about *using* modus tollens. As will be illustrated in a moment, this is trivial if prompted. But it will seem difficult if the neglect default is not jolted in a way that experience in the world teaches us when to use modus tollens. And the reason that is so, as is now on the table, is that if we did not overwhelmingly neglect opportunities to apply modus tollens we would be crippled since opportunities for that are continually at hand.[7]

If that is so, then to suppose — as discussions of Wason commonly suppose — that ordinary folk cannot easily handle modus tollens outside deontic situations (or some such special set of contexts) would be to miss the point of what makes Wason difficult. Rather, anyone of normal intelligence easily — so easily that we do not even notice we have done it — uses modus tollens when its use is prompted. But in odd or artificial or unfamiliar contexts an effective cue to offset the default may be missing. In Wason, it is apparent that subjects usually fail to notice the modus tollens inference. But if the difficulty comes from neglect defaulting, not from intrinsic difficulty in managing a modus tollens inference, then nothing more than Grice's (1989) pragmatics of language use might be needed to dispel the difficulty.

Consider this bit of dialogue. Jim and Jack are waiting for George.

JIM: (glances at watch)

JACK: If he's running late he would have called.

JIM: Ok.

You are scarcely likely to be left puzzled by this, though there is nothing here of the obligatory or rule-following (deontic) character often treated as what is needed to prompt modus tollens (Cheng & Holyoak, 1985). To make sense of the bit of dialogue you must have seen that Jim understood Jack to implicate (using the Gricean term) that in fact George had not called. And Jim must have inferred (modus tollens) that since George apparently did not call, he is apparently not running late. All that was subjectively instantaneous for Jim, and for you as well. And the 90% or so of subjects who fail to notice the modus tollens inference in the Wason test

[7] Oaksford & Chater (1998, and, in a somewhat different way, Sperber et al., 1986) propose variants turning on claims that the not-q card is seen as pragmatically inconsequential. But the arguments are tied to Wason's original formulation, which (arguably, at least) leaves a bit of ambiguity in interpretation. But there is no way that the stark formulation I will use (described later) could plausibly be misread — and by about 90% of subjects! — in a way that would warrant seeing the not-q card as pragmatically irrelevant.

could also see that. So why is modus tollens so easy here . . . indeed apparently automatic, requiring no conscious thought at all? Apparently because Jim's remark makes George's not calling (even though that is only implicit) conspicuous as something Jim thinks Jack should notice, which apparently is all that is needed to prompt a one-step modus tollens inference.[8] And that suggests that, if the neglect defaulting account I've been sketching is correct, it may be possible to cure the Wason illusion by some trivial manipulation that unobtrusively, just exploiting the pragmatics of ordinary language, prompts a one-step inference from noticing the not-2 card. This turns out to be correct.

Indeed an example has been available since very early in what is now four decades of discussion of the Wason problem. Newstead and Evans (1974) noticed that if the rule was framed as: “if A then not 3” (instead of “if A then 2”), then the difficulty largely disappears. Newstead and Evans attributed this to a tendency to mindlessly match whatever cards were mentioned in the rule. But in terms of neglect defaulting, a better explanation would be that the unusual wording of the rule alerts a person to attend to whether a “3” is seen. Given the alert provided by the odd way of framing “if A then 2” subjects see the modus tollens inference almost as easily here as in the Jim & Jack bit of dialogue.

Which suggests a cure that does not change a word of either the problem statement or the question posed. Still using “if A then 2,” suppose that instead of asking which of [A, D, 2, 3] need to be checked, subjects are asked which of [A, 3] need to be checked. Then instead of the usually predominant failure to check the “3”, subjects predominantly do check the “’3.” Indeed, consistent with the one-step aspect of neglect defaulting, there seems to be a mild tendency to miss the modus ponens inference that is otherwise rarely missed. The notorious illusion has vanished. The “flat” presentation of all the possibilities in the standard presentation of [A, D, 2, 3] does not jolt the default neglect of modus tollens. But presenting [A, 3], by the Gricean pragmatics, unobtrusively suggests alertness to A and 3, giving the not-q card (“3”) the bit of salience that George's not calling gets in the bit of dialogue. Which, as is consistent with neglect defaulting, is all that is needed to get past the usual illusion.[9]

[8] *Writing out* the modus tollens inference requires multiple steps. But writing out the inference that in a right triangle $c^2 = a^2 + b^2$ takes a much larger multiple of steps. But we don't need to run through the steps to use what we know about right triangles. From experience in the world we all know modus tollens at least as well we know canaries are birds. If canaries were as common as opportunities to apply modus tollens, it would take something to prompt us to notice a particular canary.

[9] A complaint might be that seeing [A, 3] puts the falsifying combination right in front of the chooser. But how would the chooser recognize that it is a falsifying instance without implicitly using modus tollens?

A particularly severe within-subjects test would be to prime subjects by giving the easily-solved version, immediately followed the notoriously hard-to-solve but logically identical standard version:

1. Four cards have been marked with a letter (A or D) on one side, and a number (2 or 3) on the other.

Please circle any of the two cards shown which (if turned over) might show the claim is false.

CLAIM: If the letter is A, the number must be 3.

$\boxed{A}$ $\boxed{2}$

2. Same problem but with a different claim and showing all four cards:

CLAIM: If the letter is D, the number must be 2.

$\boxed{D}$ $\boxed{A}$ $\boxed{2}$ $\boxed{3}$

And although I have not run this enough to have confident statistics, it is clear that subjects, who have included highly sophisticated game theorists and economists as well as undergraduates, very commonly get the second version wrong even though successfully primed to see the right response only a moment before with the same problem, very slightly disguised, but presented in a way intended to nudge the modus tollens switch. In another way subjects have also been primed to respond with “2,” so the effectiveness must also reflect that (so the test needs to be run also with “if D then 3”, which if the contradiction was still elicited often would be a remarkably strong result). Still, together with the 3-cards and CR games, the very simple cure of the notorious Wason illusion by the two-card choice, and the ease with which many subjects can then be coaxed away from logic they have just used, perhaps might be sufficient to jolt confidence that something as bizarre as strong neglect defaulting could be governing the intuitions of sophisticated subjects. For as you may have discovered if you were not already familiar with this material, professors as well as their students are vulnerable, even professors who have made a career of studying oddities of judgment.

2.4 The New York puzzle

And here is an even simpler example, developed from an illusory effect discussed in another context by Johnson-Laird (2005, p. 195). It involves no judgment of probability (as in 3-cards) or use of modus tollens (as in Wason) or strategic interaction (as in the CR games). Indeed nothing is required to get the answer right beyond ability to read English.

1. Just one of the following claims is true.
A) New York or Boston or both are Type X.
B) New York or Chicago or both are Type X.
Is it possible that New York is Type X?
Yes___ No___

2. One of the following claims is true, but not both:
A) New York or Boston or both are Type X.
B) New York or Chicago or both are Type X.
Is it possible that New York is Type X?
Yes___ No___

Even very sophisticated subjects tend to respond "yes" to (1) but "no" to (2) though in a completely trivial way the two questions are identical. The "one-step" response to (1), primed by "true" is that if New York was Type X that would be consistent with A and B. It takes a step further to see that would violate the premise that only one is true. The "one-step" response to (2), primed by "but not both," doesn't need the 2nd step.

2.5 The trapezoid puzzle

A simple problem from 8th grade geometry provides a yet another instructive example. See the Appendix.

3 Conclusion

And a conclusion: It would be naive to suppose that these striking effects only occur in simple puzzles. Rather, we ought to expect such effects to be significant in contexts outside the range of familiar experience with perceptible feedback, but easy to make visible only in the artificially simple contexts of simple puzzles. Beyond the range of familiar experience conditions that make the possibility of adverse neglect defaulting consequential certainly arise in cutting edge science, and even more easily in politics. What passes for "moral heuristics" that affect political judgments may sometimes conceal neglect defaults. Like the other cases I have described, people might not consider alternatives to their perception of a situation. (In other cases, heuristics might be used even when the alternatives are mentally available, and these heuristics are not what I have been describing.) Baron (in press) treats this topic in some detail. Thaler and Sunstein (2008) deal with a variety of related effects in the context of individual choices lacking feedback prompt enough to alert a chooser that they have done something that in hindsight they are likely to regret. I hope to lay out the complexities of cognitive effects in the context of public sentiment on large political choices in a forthcoming book. But the detail that is required to make a convincing case lies well beyond the scope of this note.

References

Baron, J. (in press). Cognitive biases in moral judgments that affect political behavior. *Synthese*.

Camerer, C. F., Ho, T-H., & Chong, J-K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics, 119*, 861–898.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics, 22*, 684–708.

Charness, G. & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics, 117* 817–69.

Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391–416.

Crawford, V. P., & Iriberri, N. (2007). Fatal attraction: Salience, naïveté, and sophistication in experimental "hide-and-seek" games. *American Economic Review, 97*, 1731–1750.

Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99* 3–19.

Evans, St. B. T. (1977). Lingistic factors in reasoning. *Quarterly Journal of Experimental Psychology, 29*, 297–306

Falk R. & Bar Hillel, M. (1982). Some teasers concerning conditional probabilities. *Cognition, 11*, 109–122.

Fox, C.R. & Levav, J. (2004). Partition–Edit–Count: Naive extensional reasoning in judgment of conditional probability. *Journal of Experimental Psychology: General, 133*, 626–642.

Grice, P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Kahneman, D. & Frederick, S. (2005). A model of heuristic judgment. In Holyoak, K. J. & Morrison R. G. (Eds.), *The Cambridge Handbook of Thinking and Reasoning,* pp. 267–293. New York: Cambridge University Press.

Falk, R. & Bar Hillel, M. (1982). Some teasers concerning conditional probabilities. *Cognition, 11*, 109–22.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*, 1–53.

Johnson-Laird, P. N. (2005). Mental models and thought. In Holyoak, K. J. & Morrison R. G. (Eds.) *The Cambridge Handbook of Thinking and Reasoning*, pp. 185–208. New York: Cambridge University Press.

Margolis, H. (1987). *Patterns, thinking and cognition*. Chicago: University of Chicago Press.

Margolis, H. (1996). *Dealing with risk.* Chicago: University of Chicago Press.

Margolis, H. (2007). *Cognition and extended rationality*. London: Routledge.

Oaksford, M. & Chater, N. (1998). *Rationality in an uncertain world.* Hove, UK: Psychology Press.

https://doi.org/10.1017/S1930297500000929 Published online by Cambridge University Press

Roberts, M. J. (2000). Individual differences in reasoning strategies: A problem to solve or an opportunity to seize. In W. Schaeken G. De Vooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies*, pp. 23–48. Mahwah, NJ: Erlbaum.

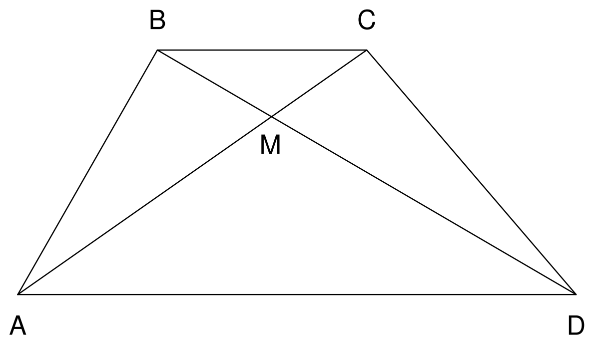
Sperber, D., Cara, F. & Girroto, V. (1986). Relevance theory explains the selection task. *Cognition, 57*, 31–95.

Thaler, R. H. & Sunstein, C. R. (2008). *Nudge.* Yale University Press.

Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of Reasoning*. Cambridge, MA: Harvard University Press.

Appendix: The trapezoid puzzle

And we can see yet another odd effect explainable by one-step defaulting in this a problem from 8th grade geometry. Very few people get it right without a lot of effort.



In the trapezoid, triangle ABM might be greater, smaller, or equal in area to triangle CDM. The question is which of the possibilities holds. And you must prove your answer is right. The solution, and an explanation of how one-step defaulting is relevant here, is in the note, but here is a hint: the proof is trivial.[10]

[10] Within the trapezoid, ΔBAD = ΔCAD (and ΔABC = ΔBCD), since they share a common base and equal altitudes. Subtract the common area ΔAMD, and we are left with ΔABM = ΔDMC. How can sophisticated subjects fail to quickly see that? Faced with this really trivial puzzle, nearly everyone responds as if blind to what is in plain sight. What is beyond the two triangles directly in play is neglected. We have a situation very like the "canary" example mentioned earlier, for the case where the 1st response is likely to be "sings" rather than "bird." The visual salience of point M in the diagram, together with the salience of "opposite angles are equal" for anyone who has had a course in geometry, make it hard to avoid a focus on the opposite angles as first response. But it leads nowhere. Only if you somehow are prompted away from that response do you easily notice the solution. I first encountered this curious problem at a dinner with Robyn Dawes.