

The Side-Effect Effect and “Intentionally”

Joshua Knobe’s (e.g. Knobe 2003, 2006) work has led many philosophers to accept the “side-effect effect” (SEE):

SEE: Foreseen, negative, but unintended outcomes are judged by the folk to be intentional actions – the folk concept of intentional action is systematically affected by normative judgments.

A key source of disagreement about SEE concerns what exactly respondents aim to convey when they agree that the CEO of Knobe’s vignettes “intentionally harmed the environment”. Where Knobe believes that they thereby attribute an intentional action of harming to the CEO, opponents of SEE argue that respondents seek to do something else, such as blame the CEO, express criticism of him, or signal that a norm violation has occurred (e.g. Adams and Steadman 2004a, 2004b, 2007; Lindauer and Southwood 2021).

The debate about this issue remains, we argue, difficult to resolve by conducting surveys in which responses to vignettes are elicited (e.g. Systma, Bishop, and Schwenkler, 2022). A method that clarifies what ordinary people seek to convey in their use of “intentionally” or “intentionally harmed” is needed.

To shed light on this, we examined “intentionally” in the Corpus of Contemporary American English (COCA, Davies 2008), where it occurs 7,511 times (ca. 7 per million). The adverb typically occurs immediately before the verbs and adjectives it modifies. We identified 955 unique verbs accounting for 4,347 occurrences and 444 unique adjectives accounting for 873 occurrences in this position. In addition, we examined a random sample of 100 concordance lines.

Our analysis suggests that “intentionally” predominantly occurs in negative contexts involving acts of violence, deception, theft, oppression, and emotional harm. For example, “intentionally misleading” is the most frequent adverb-adjective collocation. Neutral contexts are less common and often indicate design features, as in "This page is intentionally left blank." Positive contexts are rare, occurring in less than 10% of cases.

These preliminary results present, we argue, a challenge to SEE. To appreciate the challenge, note that defenders of SEE assume that “intentionally” is used by respondents to attribute intentional actions. One way to defend such a view of “intentionally” is to note that although a small group of verbs, such as “murder”, *imply* that the relevant action was performed intentionally, the vast majority of verbs have no such implication. In ordinary contexts, when one says of an agent that they did something – broke an egg, closed a door, pushed a chair aside – it will not always be clear whether the action was intentional or not. Some linguistic device(s) is needed to clarify this. It seems not unreasonable to suppose that “intentionally” is one such device.

ASSUMPTION: “Intentionally” is used to specify whether an action is intentional or not.

And yet, even if ASSUMPTION is true of *philosophers’* usage of “intentionally”, there is some reason to doubt that it is true of *folk* usage. We focus on two issues here. First, one would expect

“intentionally” to appear in positive contexts more often than it does. After all, verbs like help, assist, benefit, among others, will also sometimes need clarification concerning whether or not the relevant action is intentional. ASSUMPTION sits uncomfortably with the fact that “intentionally” is found in very few positive cases – fewer than 10%, in fact.

Second, “intentionally” was found to modify verbs such as murder, manipulate, deceive, and sabotage. Note that these verbs are not in need of clarification concerning whether they are carried out intentionally: each implies that the agent intended to do the thing in question. If ASSUMPTION were true of everyday usage, the use of “intentionally” can be redundant, and because of this, potentially confusing for an audience. On the assumption that communication avoids redundancy, is co-operative, and effective (Grice 1989), this is surprising.

These two results challenge ASSUMPTION. And because SEE relies upon it, they also challenge SEE.

These results also raise the difficult question of why so many respondents – although notably not all – agree that the CEO intentionally harmed the environment. Drawing from research in linguistics, we consider two hypotheses:

- H1: Lexical Priming
- H2: Semantic Change

H1: Corpus data suggests that "intentionally" carries a negative semantic prosody (Sinclair 1996): through habitually co-occurring with negatively charged words such as “kill”, “harm”, “misleading”, and “deceptive”, it has become associated with negative evaluation in speakers' minds. Language users are thus primed by their previous linguistic experience (Hoey 2005). Consequently, the word combination "intentionally harms" aligns with folk expectations, it is in harmony with the established semantic prosody, whereas "intentionally helps" violates these expectations and creates dissonance.

H2: Over time, habitual social and linguistic contexts of use may lead to gradual meaning change. "Intentionally" may be subject to at least two possible mechanisms (Traugott and Dasher 2002):

1. Grammaticalization – loss of lexical meaning and gain of grammatical function, here serving as an intensifier (cf. "pretty ugly" ~ "intentionally malicious")
2. Conceptual metonymization – change through habitual invited inference, here between attribution of intentionality and assignment of blame (cf. the shift from deontic to epistemic "must")

Importantly, irrespective of the specific mechanism involved, historical linguistics demonstrates that implicatures can become conventionalized, that is, encoded in the meaning of a linguistic item. We argue that H1 or H2 show promise in explaining respondents' willingness to agree that the “CEO intentionally harmed the environment” and their unwillingness to agree that the “CEO intentionally helped the environment”.

If our corpus analysis of “intentionally” is correct, there is significant pressure to reject ASSUMPTION. This gives us, then, reason to doubt the correctness of SEE. Finally, there is good

reason to look beyond current parameters of the debate about SEE – to phenomena such as semantic prosody and grammaticalization – for insight into respondents’ use of “intentionally”.

Word Count Before References: 937

References

Adams, Fred and Annie Steadman (2004a). “Intentional action in ordinary language: core concept or pragmatic understanding?” *Analysis*, 64: 173–181.

Adams, F., & Steadman, A. (2004b). “Intentional actions and moral considerations: Still pragmatic”, *Analysis*, 64, 264–267.

Adams, Fred and Annie Steadman (2007). “Folk concepts, surveys, and intentional action.” In C. Lumer and S. Nannini (eds.), *Intentionality, Deliberation, and Autonomy: The Action-Theoretic Basis of Practical Philosophy*, pp. 17–33, Aldershot: Ashgate Publishers.

Grice, P. (1989) *Studies in the Way of Words*. Harvard.

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. Routledge.

Knobe, J. (2003). “Intentional action and side effects in ordinary language”, *Analysis*, 63, 190–193.

Knobe, J. (2006). “The concept of intentional action: A case study in the uses of folk psychology.”, *Philosophical Studies*, 130, 203–231.

Lindauer, M. and Southwood, N. (2021). “How to Cancel the Knobe Effect: The Role of Sufficiently Strong Moral Censure.” *American Philosophical Quarterly*, 58(2): 181–186.

Lumer and S. Nannini (eds.), *Intentionality, Deliberation, and Autonomy: The Action-Theoretic Basis of Practical Philosophy*, pp. 17–33, Ashgate Publishers.

Sinclair, J. (1996). “The search for units of meaning.” *Textus*, 9, 75-106.

Systema, J., Bishop, R., and Schwenkler, J. (2022) “Has the Side-effect effect been cancelled? (No, not yet.)”, *Synthese*, 200: 395, 394-395.

Traugott, E. C. and R. B. Dasher (2002). *Regularity in Semantic Change*. Cambridge University Press.

Davies, M. (2008–). *The Corpus of Contemporary American English (COCA)*. <https://www.english-corpora.org/coca/>