Comment

Targeting Risks

Paul Slovic and Baruch Fischhoff

A fundamental premise of government and industry safety policies is that it is possible to reduce the rate and severity of accidents by improving the design of machines and the environments in which they are used and increasing the skill of their human operators. Wilde's Theory of Risk Homeostasis constitutes a frontal attack on this premise and, accordingly, on the safety measures based on it. Wilde is to be congratulated for carefully and explicitly setting out a fascinating theory of risk-taking behavior. In addition to its theoretical interest, his proposal has important practical implications making it worthy of detailed, critical examination. To the extent that it is valid, the theory points to a need for significant changes in how society allocates its resources so as to maximize health and safety without sacrificing economic vitality and productivity. Our comments here shall focus on three general topics. The first is the nature of the theory and its assumptions. The second is the potential for testing the theory. The third is the inherent limitations of theories that attempt to deal with risk issues in isolation. Although we remain skeptical of the truth of this theory, we find it highly productive of research questions, answers to which could markedly improve safety policy.

1. THEORY AND ASSUMPTIONS

1.1. History

Homeostatic theories have been around for many years. In 1859, Bernard postulated a general law of constancy of the internal physiological environment to account for the regulatory mechanisms that held the blood's concentration of glucose at a nearly constant level. Observing a variety of such steady physiological states, achieved by interacting mechanisms, Cannon coined the term "homeostasis." He found it to be a useful principle for integrating numerous physiological phenomena that had appeared either paradoxical or contradictory. Within the behavioral sciences, Dempsey argued that many intellectual and social activities appeared directed toward homeostasis, but had difficulty achieving it. Noting the wide swings of social welfare embodied in war and peace, hunger and abundance, employment and idleness, he observed:

It is perhaps not surprising, in view of the recent evolutionary emergence of the mind as a homeostatic instrument, that in some respects its functions should be less efficient than the older physiological ones. In such respects the body physiologic has evolved methods of operation better than those thus far prevailing in the body politic. (pp. 233-234)

Although provocative and engaging, homeostatic theories have drawn criticism directed at their validity and usefulness. Wilde's theory is a novel application to health and safety of the homeostasis concept. In considering it, we will avail ourselves of some of the other comments that have emerged during homeostasis' long and controversial history. We note in passing that a more systematic examination of homeostatic theories of social processes could prove instructive.

1.2. Where Might the Theory Go Wrong?

The core of Wilde's theory is the hypothesis that people have a target level of risk in different activities (not necessarily the same target for all activities).
They are assumed to adjust their behavior so as to achieve that level of risk. As a result, safety measures that fail to reduce the target level will fail to reduce risk, because people will engage in actions that will offset any safety gains those interventions might have achieved. Thus, measures such as mandating the use of seat belts or improving vehicle design should, according to the theory, have only temporary effects on motor-vehicle accident rates.

Note that this theory does not claim that all safety measures will be ineffective, only those in which the current risk level is the desired one. Wilde uses the term “steady-state error” to describe situations in which current risk levels are undesirably high. From this perspective, one way to improve safety is to change people’s perceptions of the current risk, so that it is seen as above their target level. For example, people appear to underestimate the risks from motor vehicles. When informed that the probability of one or more serious injuries across a lifetime of driving is on the order of 0.33, they seem much more willing to wear seat belts.\(^5\) (A second safety strategy coming from the theory would be to alter the target level of risk, a topic to which we return later.)

Although the homeostatic theory might be interpreted as showing that “safety measures are futile,” it points to a variety of situations in which such measures may actually be quite effective. For example:

1.2.1. When Current Risk Levels Are Unacceptable

As discussed immediately above, when current risk levels are not acceptable, people should be quite eager to adopt safety measures rather than engage in “compensatory” actions. Understanding people’s beliefs about current levels of risks and their views about the appropriateness of these levels would, according to the theory, be a precondition for launching any safety program.

1.2.2. When Habits Are Strong

Many behaviors are strongly conditioned by habit, so much so that people could not change them even if they wished to. Driving habits would seem a likely candidate for such immutability. It is hard to imagine experienced drivers adjusting their style in response to interventions such as padded dashboards or stronger highway median barriers.

1.2.3. When No Direct Compensation Mechanism Exists

Padded dashboards and better highway barriers are also examples of safety measures for which there are no directly applicable compensatory mechanisms. One can always act more recklessly, but not in a way that restores dashboard impact to its prior level or maintains the probability of being hit by an errant car going the opposite direction (or being that car). Although compensation through other means (e.g., driving while more tired, removing one hand from the steering wheel) is conceivable, linking diverse safety measures and responses so that they balance one another requires quite a complex and sophisticated cognitive model. Inserting such a model into the theory of risk homeostasis might make the theory itself less credible.

1.2.4. When the Risk Reduction Is Underestimated

In order to respond to the reduction in safety, people must be able to perceive it. If they underestimate the improvement, then their adjustment will be “too small” leaving a residual increase in safety. Conversely, a safety measure whose impact is exaggerated would lead to a net reduction in safety—all this assuming that the theory is true. Behavioral studies show a variety of limits to the validity of people’s risk perceptions.\(^7\) Some limitations are due to the quality of information people receive; others follow from what they do with that information. All would loosen the linkages proposed by the theory. Perhaps the most common form of misperception is exaggerating one’s own skill and safety (see, for example, ref. 9). If people already believe their personal risk to be extremely small, they may fail to recognize the risk reduction provided by a new safety device, hence, they may not act in ways that counteract their new safety gains.

1.2.5. When the Safety Measure Is Not Visible

The extreme case of underestimating a measure’s effectiveness is not to know that it exists. Some safety measures are effectively invisible, which should make them immune to countervailing homeostatic measures. For example, few drivers know enough about automotive engineering to track improvements in frame construction.
These five conditions are ones in which adaptive processes could not work even if the theory of risk homeostasis were generally true. Whether they are the exception or the rule in human experience requires a more detailed analysis than is possible in this comment. If they are common, which would be in keeping with Dempsey's claim that sociopolitical systems have difficulty regulating themselves, then the effects of homeostatic tendencies would be negligible or erratic.

These conditions are, moreover, but a subset of those in which homeostatic mechanisms are known to fail for behavioral systems. Three more examples will illustrate the kinds of processes that need to be considered:

1.2.6. Adaptation

Psychological and social systems, like biological ones, tend to respond less and less to repeated stimulation of a given type. In time, the effect of any stimulus may become completely neutralized because the organism has brought its "adaptation level" in line with the level of stimulation. For example, the eye accommodates to wide variations in lighting and the ear eventually disregards background noises that initially were quite disturbing. Theories of adaptation assert that it produces new states of equilibrium from which behavior can be measured, predicted, and understood, without implying that the goal of behavior is a state of equilibrium.(10) Adaptation is a powerful and prevalent process. It may, in many circumstances, lead people to accommodate to and accept reduced levels of risk rather than seek to restore the prior risk level. Indeed, anecdotal and survey evidence suggests that people adapt quickly to increased health and safety, coming to view it as normal and as their right, perhaps even desiring more of it. At a time when health and longevity are at their highest levels, U.S. society appears to be more concerned about risk reduction than ever before.(11,12)

1.2.7. Need for Variety

Some theorists have postulated that people have a "need for variety,"(13) which would be the antithesis of homeostasis. According to this view, novelty, change, and complexity are pursued because they are inherently satisfying. The search for variety might lead people to seek different levels of risk in different realms of their lives (home, work, leisure). People might also vary the risk level within a realm for the sake of stimulation, experience, or learning what it is like.

1.2.8. Locus of Stability

Homeostatic theories depict a flexible, even inventive organism maintaining its equilibrium in the face of a changing environment. A contrasting view, increasing in popularity among motivation theorists, is that the constancy of observed behavior is actually a consequence of the stability of the controlling environmental conditions.(14) Following this view, risk levels would be expected to be uniform only when behaviorally significant features of the risk environment remain stable.

To summarize, the theory of risk homeostasis is one exemplar of a large and respected category of theories for describing behaviors that seem to maintain an equilibrium condition. Unlike the others, it postulates a volitional mechanism for achieving that state, the desire to maintain a constant level of risk. Like the others, it has inherent strengths and weaknesses. Among the former, are its generality and simplicity. Among the latter, are the existence of situations in which it could not operate and the existence of known behavioral effects that contradict it. If those situations are common and those effects are powerful, then the practical implications of the theory might be quite limited—even if it is true. The status of its empirical validation is the topic of the next section.

2. TESTING THE HOMEOSTATIC THEORY OF RISK

Wilde carefully pointed out that he has not tested his theory. Rather, he was advancing it as a hypothesis, along with some data that might support or elaborate it. Clearly, the homeostatic theory needs to be verified if it is to be taken seriously as a model of behavior and a guide to policy.

Such verification appears, however, to be difficult. The theory contains numerous unspecified parameters that enable it to explain data post hoc without enhancing its predictive capabilities. Each of the limiting conditions noted in the previous section provides a potential reason (or excuse) for why
homeostasis failed to express itself in a particular situation. Thus, if a safety intervention reduced accident rates (without obviously increasing motivation to be safer), one could defend homeostasis by claiming that people had misperceived (underestimated) the intervention's effectiveness, or that there had been insufficient time for adjustment, or that risks were previously above people's target level for that activity. Like any other theory, homeostasis can also be defended against apparently contradictory evidence by attacking the research design. For example, one might claim that the spatial frame of observation was too narrow to show that the risk reductions in a specific location were actually offset by an increase in risk across a wider geographic area. It may be particularly easy to generate claims based on imperfect measurement of the target level, whose labile nature, with sensitivity to gain and losses and other momentary influences, may make it quite elusive. Defended in these ways, the theory does not appear to be falsifiable.

O'Neill(15) describes four types of evidence that can be brought to bear on the theory. These are (a) physiological measures of risk,(16) (b) verbal ratings of risk across situations,(17) (c) longitudinal comparison of accidents or risk taking before and after a safety change, and (d) contemporaneous comparison of accidents or risk taking in populations with differing environmental safety levels.

O'Neill goes on to show the methodological difficulties associated with each type of study. Physiological responses are subject to arousal by factors other than risk. Ratings respond to a variety of contextual factors that complicate comparisons across contexts.(18) For example, people may recalculate their rating scores so that the extreme ratings correspond to the extremes of the stimuli in each context.(19) Thus, people might rate quite different behaviors as equivalent in risk, not because they had been seeking a constant risk level in different situations, but because of the way they use rating scales. Although it is naturally appealing to compare accidents and risk taking before and after safety interventions, those comparisons are only useful if one can identify and control (either statistically or experimentally) effects due to extraneous changes in the environment. Such effects can either produce spurious changes in safety levels or frustrate potentially effective interventions.

In light of these many possible pitfalls and the important consequences of this research for safety policy (which makes studies subject to sharp criticism), it is not surprising that the research to date has been equivocal. For example, Peltzman(20) used regression analysis to show that motor vehicle safety regulations had not reduced overall accident rates. His work was attacked by Robertson(21) on various grounds, including: (a) Peltzman's model did not predict fatality rates accurately prior to regulation; (b) deaths involving cars subject to regulation were not separated from deaths involving vehicles not subject to regulation; and (c) alcohol consumption and youth involvement in crashes were not measured properly. Attempting to improve on the analysis, Robertson found that fatality rates during the regulatory period from 1966-1972 were well below the rates projected on the basis of rates from the preregulatory period. He concluded that there was no evidence of increased "risky driving" in regulated vehicles. Joksch(22) also criticized Peltzman's findings for many of the same reasons cited by Robertson. These criticisms have been rebutted by Peltzman(23) and the rebuttal rebutted by Robertson.(24)

Lindgren and Stuart(25) applied regression analysis to the effects of traffic safety regulations in Sweden from 1965-1973. They estimated that these regulations were accompanied by a substantial reduction in the fatality rate for vehicle occupants and a small (although statistically nonsignificant) reduction in the fatality rate for nonoccupants. Assuming that Lindgren and Stuart's methods were acceptable, the proponent of homeostasis could still argue that the Swedish regulations included strict speed limits, which curtailed the opportunity for risk substitution. That argument would preserve the theory's validity by restricting its realm of applicability (i.e., admitting that enforcing speed limits was sufficient to block the homeostatic processes).

Tests of the theory in the real world are inherently difficult because safety measures often are not implemented in the sort of decisive, consistent manner that would allow clear tests of their efficacy. Such strict enforcement as accompanied the Swedish speed limits may be the exception rather than the rule. Failure to consider noncompliance can blur the effects of safety measures. For example, studies of aggregate industry-wide accident data following the setting of safety standards by OSHA found no perceptible reduction in injury rate.(26-28) However, when examining the subset of plants where OSHA regulations were strictly enforced, Cooke and Gantschi(29) found sizable reductions in days lost due to injury. Plants with their own voluntary safety programs (jointly administered with unions) were also successful in reducing lost work days.
Incomplete compliance obscures effects by pooling individuals who have received the “treatment” implied by the measure with those who have not. It can also lead to a mixture of compliers and noncompliers whose mutually inconsistent behavior interacts to produce new risks. For example, Wilde cites as evidence of homeostasis the constant accident rate that accompanied a program in Holland that raised from 37 to 80 the percentage of drivers who used low-beam headlights after dusk instead of parking lights (the usual practice at the time). Prior to the intervention, cars with low-beam headlights turned on had been found to be less often involved in accidents than cars with only parking lights on, so it was expected that increasing the percentage of headlight users would increase the overall safety level. However, the 20% of drivers who failed to comply became a great threat to their fellow drivers who had become accustomed to seeing headlamps. A better test of the theory would have been to examine accident rates following mandatory use of low-beams by all drivers.

Many of the problems that arise here are common to all attempts to study interventions into people’s behavior. For example, over the last half century, clinical psychologists have developed a sophisticated methodology for assessing the effects of their treatment programs.\(^1\) One standard procedure is to measure the fidelity with which a treatment has been applied. Incomplete application can be used to argue that a program has not been given a fair test; it may also mean that the program cannot be applied in realistic conditions.

Another question that has worried clinical psychologists is how to measure the mental health states that constitute the outcome of treatment (or non-treatment). The difficulty of measuring what really interests them (mental health) has led to the development of intermediate criteria that are expected to be associated with mental health (or the lack of it). The difficulty of assessing the rates of serious accidents (which are, fortunately, quite unlikely events) and linking them to interventions has also led to the use of intermediate criteria in the accident field. For example, Evans, Wasielewski, and von Buseck\(^3\) observed the difference in following headway\(^2\) in freeway traffic for seat-belt users and nonusers in two communities, one of which required seat-belt usage and one of which did not. Although the incidence of seat-belt use differed greatly in the two communities, there was no evidence of danger compensation. In both locales, seat-belt users allowed greater headway than nonusers. As Wilde notes, however, the value of such studies depends on the strength of the link between the intermediate variable (wearing seat belts) and the ultimate one (safety).

Furthermore, as with all field studies, the interpretation of the research by Evans et al. depends upon the comparability of the people in the different conditions. Using Evans et al. as evidence against homeostasis requires one to assume that the compelled users are otherwise similar to nonusers. The very fact that their community enacted such a law might, however, suggest that its citizens are particularly cautious.\(^3\) The obvious alternative to field studies is experimental research, which provides more comparability at the cost of some degree of realism. Thus far, there have been few studies of this kind. One intriguing possibility would be to have drivers with and without safety devices (seat belts, stronger vehicle frames, dual brakes, etc.) tour a standard course in a vehicle instrumented to record various performance measures.

### 3. DECISIONS ABOUT RISKS

A distinctive feature of the principle of risk homeostasis is that it is expressed in terms of risk alone. People acting in accordance with it would make choices between alternative modes of behavior solely in terms of the level of risk that each entails—choosing the action whose risks are closest to their target level. In doing so, they would not take account of the other consequences (e.g., benefits and nonrisk costs) associated with those options. Taken literally, this would mean holding resolutely to that level even if there were considerable benefits to be obtained by tolerating a bit more risk or if considerable additional safety could be gained at relatively little cost.

If one believes that risky decisions are not just concerned with risks, then radically different inter-

---

\(^2\)Following headway is the time interval between a vehicle and the vehicle immediately ahead in the same lane arriving at the same point on the roadway. It was selected as a measure of driving intensity and an indicator of driver risk taking.

\(^3\)A recent study by Geller\(^12\) was designed to overcome the comparability problem. Geller took advantage of a factory campaign that provided incentives for seat-belt use to study the behavior of the same drivers when belted and when not belted. Measurement of driving speed on a curvy and narrow two-lane road in the absence of other traffic showed no influence due to seat-belt usage.
pretations arise for evidence that has been cited demonstrating homeostasis. For example, increasing the stability of tractors should reduce the associated accident rate providing that their usage remains otherwise unchanged. Farmers may, however, realize that they can use the new tractors to plow steeper hillsides, although doing so will increase their risk level (perhaps even to where it was with the old tractors). In return for that risk, farmers are getting a substantial benefit, the yield from their newly arable land. For the farmers, the safety intervention that increased the tractors' stability is a success, giving them increased benefit at no increase in risk (ignoring, for the moment, any increase in tractor costs). Society, too, should be better off for the increase in productivity.

The intervention would be a failure only if one looked at safety in isolation, as might the safety officials who mandated the change—and whose performance was judged solely on the basis of accident statistics. If policy makers are placed in a position where they can consider only one consequence of their actions, then there is something wrong with our social and institutional arrangements. Regulatory decisions that consider only safety are as inadequate as corporate decisions that consider only the quarterly bottom line.

The same kind of narrowness leads to attempts to define what characterizes an acceptable level of risk. When people make risky decisions, they choose options, one of whose consequences is some level of risk. One cannot infer from their decision that they are satisfied with that level of risk or that they would not "accept" higher risks if another option came along that offered considerably more benefit in return for a modest increase in risk. (33)

Accommodating non-risk consequences calls for a rather different kind of theory than risk homeostasis. One alternative is provided by O'Neill, (15) who developed a decision theoretic model of risk compensation. O'Neill argued that drivers rationally attempting to maximize the expected benefits of driving minus the expected losses due to an accident may, under certain circumstances, respond to a safety improvement by increasing their speed to the point that their accident probability rises. This behavior is not the result of compensation but of a desire to maximize gain rather than maintain a target level of risk. Stable risk levels will be observed only when there is something to be gained by accepting more risk. From this perspective, improved safety is most likely to be obtained by measures that offer no opportunities for compensating gain.

Indeed, from this perspective, the usefulness of compensation as a concept becomes doubtful. O'Neill's model is a special case of a broad class of theories known variously as expectancy theories (34) or expected utility theories. (35) These theories assume that behavior is governed by the desire to maximize some function of the difference between the gains and losses expected to result from one's actions. They are, in a sense, psychological cost/benefit theories. None of these theories has given salience to a target level of risk. The level of risk has significance only in comparison to the level of benefit.

Such theories have been the subject of extensive empirical investigation during the past quarter century. (36-39) Although they have proven to be imprecise in some respects, they do predict certain classes of behavior reasonably well. (34,40,41) Of the hundreds of studies of risk-taking behavior designed to test expectancy theories, few have observed a tendency to seek and maintain some target level of risk. (4) Whether this is due to the particular tasks studied or to the absence of target levels is a question worth investigating.

Wilde does allow for nonrisk factors, but through an indirect mechanism, changes in the target level of risk (see his Fig. 9). For example, a driver in a hurry would be expected to have a higher target level of risk because of the greater perceived benefit of risky behaviors such as speeding or driving through amber traffic lights. Wilde's target level is thus seen to be quite labile, varying between and within individuals even from moment to moment. However, the notion of fluctuating target levels vitiates the usefulness of that concept, which should be a behavioral constant, governing actions in a wide variety of situations.

4. CONCLUSION

Wilde's theory might be interpreted as showing that safety measures are impotent, whether promulgated by industry, government, or citizens groups. We believe that, at present, there is little empirical support for the theory. There are few directly pertinent studies and these seem to be equivocal. As a result, the Theory of Risk Homeostasis should be treated as just what Wilde asserts it to be, an intri-

4A notable exception is the work of Coombs. (42)
Targeting Risks

guing set of hypotheses awaiting empirical affirmation or disaffirmation.

We wholeheartedly agree with Wilde that such empirical analysis should stand high on the agenda of those concerned with safety policy. The importance of the theory comes not just from its potential policy implications, but also from the rich theoretical apparatus it provides. Developing the theory has enabled Wilde to generate new research questions and ingenious speculations about how existing studies might be interpreted. Wilde's paper points to the complexity of human behavior in sociotechnical systems and to the naivete of some interventions—even if the theory is not true.

Should the theory eventually be supported by evidence, even then it could not be taken as showing the uselessness of safety interventions. Rather, it would point the way to more effective safety interventions. As described in the first section of this paper, there are conditions in which homeostatic mechanisms are unlikely to operate and, therefore, are unlikely to frustrate safety measures. Furthermore, the theory highlights a potentially important class of interventions—actions that reduce people's target level of risk. If effective, such actions could have far-reaching effects on the improvement of safety. Reducing people's tolerance for risk should have a salutary effect on safety even if the theory is not true. It is surprising that there have been so few attempts to pursue this strategy.

Finally, consideration of the theory of risk homeostasis points to the limitations of attempting to deal with risk issues in isolation. Just as it is meaningless to talk about acceptable levels of risk without considering the other costs and benefits that are incurred with a particular action, so is it problematic to think about people maintaining a target level of risk, oblivious to the costs and benefits associated with more and less risky behavior. The fact that society has institutions whose main charge is risk management does not ensure that people have an equally narrow focus.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation under Grant PRA-8116925 to Perceptronics, Inc. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. We have benefited from discussions of this topic with Frank McKenna and Ola Svenson.

REFERENCES

12. A. Wildavsky, No risk is the highest risk of all, American Scientist 67, 32–37 (1979).