Safety Science 109 (2018) 377–381

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

On safety, protection, and underweighting of rare events☆

D. Cohen[a,*], I. Erev[b,c]

[a] Max Wertheimer Minerva Center for Cognitive Studies, Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Israel
[b] Max Wertheimer Minerva Center for Cognitive Studies, Faculty of Industrial Engineering and Management, and the Transportation Research Institute, Technion – Israel Institute of Technology, Israel
[c] Warwick Business School, Warwick University, United Kingdom

ARTICLE INFO

Keywords:
Forgetting to be afraid
Safety technologies
Risk compensation
The peak-end rule

ABSTRACT

The current research clarifies the conditions under which safety enhancing interventions backfire. A laboratory experiment compares three repeated choice conditions. In Condition Baseline, the participants were asked to choose between a safe prospect, and a counterproductive risky prospect that led to a gain, moderate loss, or a large but rare loss. The other conditions simulate safety interventions that modify Condition Baseline by protecting the participants from one of the two losses, while keeping the risky choice equally counterproductive. Results show that protection against the rare loss was effective, but the protection against the moderate loss impaired participants' earnings. The results are captured with a simple model that assumes reliance on small samples of past experiences. Implications are discussed.

1. Introduction

Analysis of fatal road accidents in the USA highlights the possibility of a disturbing trend change. The results (cf. Left-hand side of Fig. 1) show a monotonic decrease in traffic-related death rates per 1,000,000 inhabitants in the USA, from 1994 (157.0) to 2011 (104.2), and a slight increase after 2011. The death rate in 2015 was 110.4, and the death rate in 2016 was even higher: 116.1 (FARS; Fatality Analysis Reporting System, National Highway Traffic Safety Administration (NHTSA), 2017a,b). Interestingly, much of the increase has been among outside-car fatalities (i.e., pedestrians, bicyclists, motorcyclists), as opposed to inside-car fatalities (i.e., drivers, passengers; cf. right-hand side of Fig. 1).

One natural explanation for the increase in the death rate, described in Fig. 1, suggests that it reflects the negative side effects of new technologies, such as smartphones, that increase the benefits from risky behavior. For example, it is possible that in certain cases texting while driving maximizes the driver's expected utility even when it increases the probability of an accident (Caird et al., 2014; Cook and Jones, 2011; Klauer et al., 2014; Young and Salmon 2012). This "rational cost-benefit explanation" suggests an easy solution to the apparent problem: Reducing the benefit of new reckless behaviors with improved law enforcements and bigger punishments.

The current paper examines the feasibility of a second contributor to the pattern summarized in Fig. 1. It considers the possibility that part of the problem reflects the impact of new technologies that are designed to enhance safety, but backfire (see Dekker, 2014; Hedlund, 2000; Larsson et al., 2010; Leveson, 2011; Noland, 2013; OECD, 1990; Summala, 1996; Wang et al., 2013). Specifically, partial protection can lead people to behave as if they "forget to be afraid" (Baker et al., 2007; Reason, 1998). We hypothesize that in certain situations the impact of forgetting to be afraid can be bigger than the impact expected under risk homeostasis or risk compensation (Adams, 1995; Wilde, 1982; Hedlund, 2000). Risk homeostasis and risk compensation suggest that new safety interventions lead people to take more risks to maintain the pre-intervention risk level, and for that reason, the new interventions have a limited effect. The current hypothesis is more extreme on the one hand, and more specific on the other. It is more extreme as it suggests that in certain settings partial protection is expected to hurt the protected individuals. It is more specific, as we do not assume that the negative effect is general. Our main goal is to clarify the conditions under which partial protection backfires.

Our analysis distinguishes between two classes of partial protection that can lead people to "forget to be afraid". The first involves partial protection that improves the worst-case outcomes. Possible candidates for technologies that provide partial protection of this type include helmets, airbags, and Advanced Driver Assistance Systems (ADAS) that slow down the car when it detects a significant risk of high-speed

☆ This research was supported by the I-CORE program of the Planning and Budgeting Committee and the Israel Science Foundation (grant 1821/12). We thank Zillah Cohen and Alex Volchek for their help in proof reading the manuscript.
* Corresponding author at: Max Wertheimer Minerva Center for Cognitive Studies, Faculty of Industrial Engineering and Management, Technion, Haifa 3200003, Israel.
E-mail address: dcohen@campus.technion.ac.il (D. Cohen).

https://doi.org/10.1016/j.ssci.2018.06.020
Received 7 November 2017; Received in revised form 23 May 2018; Accepted 20 June 2018
0925-7535/

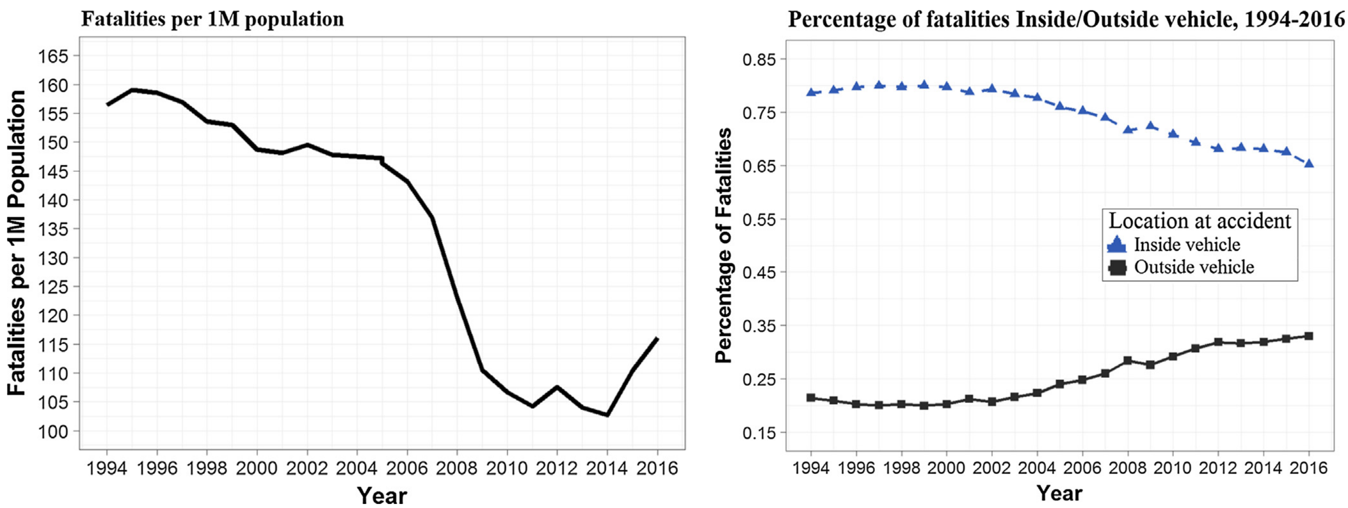

Fig. 1. Analysis of fatal road accidents in the USA, 1994–2016. Left-hand side: Fatality rates per 1 million inhabitants in the USA, 1994–2016. Right-hand side: Percentage of fatalities inside (car occupants) vs. outside vehicle, USA, 1994–2015 (NHTSA, 2017a,b).

collision. The second class involves partial protection from mild but frequent losses. Examples of technology that provide partial protection of this type, and imply a typically forgiving setting, include shock absorbers and ADAS that help prevent drivers from bumping into the car in front at low speed and stop-and-go traffic situations.

Basic research in psychology and behavioral economics suggests that both subclasses of partial protections can backfire. Improving the worst-case outcomes is likely to backfire when behavior is driven by the Peak-end rule (Fredrickson and Kahneman, 1993). The Peak-end rule suggests a tendency to remember the peak (extreme) and the end (final) experiences while underweighting the other experiences. For example, in a study by Fredrickson and Kahneman (1993), participants were asked to provide online (real-time) evaluations and a global evaluation of aversive film clips. Each clip had a long and short version. The correlation between the peak of the online evaluations and the global evaluation was 0.76. The correlation between the global evaluation and the clip duration (a proxy of the objective pain) was only 0.25. If people avoid texting while driving because of one memorable costly experience such as a rear-end collision with a car that stopped at the side of the road, technology (such as ADAS) that reduces the cost or probability of accidents of this type might increase texting while driving.

Preventing mild and frequent losses from risky choice is likely to backfire when people underweight rare events (see Barron and Erev, 2003; Etzioni et al., 2017; Hertwig and Erev, 2009). A bias toward underweighting of rare events tends to emerge with accumulation of experience. For example, the participants in Etzioni et al. (2017) were asked to control the speed of a virtual car in a simplified simulator. Speeding up increased a basic frequent gain (from 2.5 to 4.5 points), but it also increased the probability of a rare but large loss (100 points). The optimal speed was 90 km/h. The participants' average speed was higher: 97 km/h and 102 km/h in the first and the second trip, implying participants were much more sensitive to the frequent gain than to the possibility of experiencing a large, rare loss. Comparison of alternative explanations of this bias demonstrates the descriptive value of models that assume reliance on small samples of past experiences (see Hertwig et al., 2004; Erev and Roth, 2014). Reliance on small samples implies underweighting of rare events because rare events are underrepresented in a small sample.[1] For example, if decision makers rely on a sample of size 5, an event that occurs in 5% of the cases will be included in only 22% of the samples. If people avoid texting while driving due to frequent "close-call" experiences such as bumping into the car in front at low speed or stop-and-go traffic situations, technologies that reduce the risk of this cost might increase the rate of texting while driving.

2. Experiment

It is very difficult, if not impossible, to distinguish between risky behavior that reflects a rational cost-benefit analysis and reckless behavior that reflects irrational "forgetting to be afraid" biases, based on field data. Almost any risky behavior can be explained as the product of rational considerations given certain assumptions concerning the underlying utilities. To address this difficulty, the current analysis uses a simple experimental task in which we determine the subjects' incentives and ensure that the risky choice is counterproductive.[2]

The current experiment examined the numerical examples (problems) presented in Table 1 using the experimental paradigm presented in Fig. 2. In the experiment described below, each participant faced each of the three problems for 100 trials and was paid for one randomly selected trial with a conversion rate of 1 Shekel (about \$0.25) for 2 points (see *Method* section below for more details regarding participants and procedure). In each trial, the participant's task was to make a choice between the two possible options.

In problem *Baseline*, the safe option abstracts the choice to comply with a safety norm or rule. This option leads to a payoff of 0 in most cases but can also lead to a large gain (+40 with probability .05).[3] The risky option abstracts a violation of some safety rule or norm, such as texting while driving, and can lead participants to experience a major accident (−40 with a probability of .1), a minor accident (−1 with p = .5), or a gain of +2 otherwise. The *Baseline* problem represents the environment prior to new safety technology implementation.

Problem *Frequently-Forgiving* abstracts the addition (to the baseline) of protecting technology that eliminates the risks of frequent, minor losses. One example for such technology is a device that prevents minor accidents at low speeds. Problem *Better-Worst-Case* abstracts an

[1] The probability that an event that occurs with probability p is included, at least once, in a sample of size m is $1 - (1 - p)^m$. When $p < .5$, the event is included in less than half the samples of size. $m < \mathrm{Log}(1/2)/\mathrm{Log}(1 - p)$.

[2] Hedlund (2000) questions the value of laboratory experiments in the context of reckless behaviors. Part of his critique is based on the observation that most experiments (reviewed by him) only show that behavior responds to incentives, and "this is hardly news" (page 86). We believe that the current design can lead to news: Controlling the incentive structure can clarify the conditions under which partial protection backfires.

[3] The rare gain was introduced to abstract environments in which safe and relaxed driving (without reading and replying to text messages) increases the probability of a random thought that will lead to an insight. In addition, it helps increase the EV advantage of the safe prospect without using very large losses.