TOWARDS A THEORY FOR UNINTENDED CONSEQUENCES IN ENGINEERING DESIGN

Walsh, Hannah (1); Dong, Andy (2); Tumer, Irem (1)

1: Oregon State University; 2: University of Sydney

ABSTRACT

Conventional failure analysis ignores a growing challenge in the responsible implementation of novel technologies into engineered systems - unintended consequences, which impact the engineered system itself and other systems including social and environmental systems. In this paper, a theory for unintended consequences is developed. The paper proposes a new definition of unintended consequences as behaviors that are not intentionally designed-into an engineered system yet occur even when a system is operating nominally, that is, not in a failure state as conventionally understood. It is argued that the primary cause for this difference is the bounded rationality of human designers. The formation of unintended consequences is modeled with system dynamics, using a specific real-world example, and bifurcation analysis. The paper develops propositions to guide research in the development of new design methods that could mitigate or control the occurrence and impact of unintended consequences. The end goal of the research is to create a new class of failure analysis tools to manage unintended consequences responsibly to facilitate engineering design for a more sustainable future.

Keywords: Unintended consequences, Design theory, Decision making, Sustainability, Social responsibility

Contact:
Walsh, Hannah Scharline
Oregon State University
School of Mechanical, Industrial, and Manufacturing Engineering
United States of America
walshh@oregonstate.edu

Cite this article: Walsh, H., Dong, A., Tumer, I. (2019) ‘Towards a Theory for Unintended Consequences in Engineering Design’, in *Proceedings of the 22nd International Conference on Engineering Design (ICED19),* Delft, The Netherlands, 5-8 August 2019. DOI:10.1017/dsi.2019.348

1 INTRODUCTION

In April 2010, Deepwater Horizon, one of the most technologically advanced oil rigs of its time, spilled an estimated 4.9 million barrels of oil into the Gulf of Mexico, killing millions of birds and marine animals, not to mention the tragic loss of human life (Davenport, 2018). The accident led to BP paying $18.7 billion to federal and state governments – the largest environmental settlement in the United States to date (Davenport, 2018). What is known as one of the largest environmental disasters of recent memory was the result of design flaws and insufficient maintenance, which resulted in the failure of multiple safety mechanisms that were designed to prevent such an accident (Mullins, 2010).

Consider a second scenario. The internal combustion engine was first introduced in the late 1700s and would later go on to revolutionize travel, particularly when used in motor vehicles a century later. However, the designers of the internal combustion engine, while they were aware of the carbon dioxide emissions of the machine, did not predict the interaction between these emissions and the atmosphere, which generates a greenhouse effect in the atmosphere and contributes to global warming.

While both scenarios resulted in a negative impact on the environment, there is a fundamental difference between them. The first is a classic case of system failure. The second is a different kind of failure: a failure to adequately consider the broader impact of a novel technology. In 2019, novel technology-enabled products are exciting and promising – autonomous vehicles promise greater safety on the roads, electric aircraft are expected to reduce the environmental impact of air travel, and advances in robotics suggest new possibilities for search and rescue and other life saving operations. However, given their novelty, there is little historical data to help designers understand the potential risks, and it is difficult to anticipate the consequences of these technologies. Social media and recent news about Facebook in particular is an apt example of the so-called *unintended consequences* that can result from novel products.

The Law of Unintended Consequences states that for every action there is an unintended or unanticipated outcome (Ring and Madni, 2005; Schweber, 2011; Dickie *et al.*, 2016). It is commonly referenced to explain why actions can have non-intuitive consequences. The rapid progression of technology combined with their implementation into novel products and services that play a significant role in peoples' lives underscores the necessity of minimizing potential unintended consequences, if engineering design aims to be responsible. Unfortunately, the common wisdom is that nothing can be done about so-called "unknown unknowns" – that is, unidentified risks (Kim, 2012) – and organizational risk identification strategies often fail to, in the words of Mark Zuckerberg as he testified before Congress last year, "take a broad enough view of our responsibilities" (Rocha *et al.*, 2018). For our social, financial, and ethical responsibilities to keep up with the introduction of novel products and services, we need a new approach to manage unintended consequences.

To address this need, this paper develops a theory for unintended consequences in engineering design. This theory refines the Law of Unintended Consequences to be applicable to engineering design. Unintended consequences may be anticipated but ignored due to willful ignorance or traded off against other design decisions. Or, unintended consequences may not be anticipated due to the bounded rationality of engineers with imperfect information about the system they are designing and with limited capacity and time to make design decisions. The resulting behavior that is not intentionally designed-in to the system is what the authors define as an unintended consequence.

The next section presents two axioms used as foundations to develop the theory of unintended consequences. Next, in the third section, an explanatory model of unintended consequences is presented and a definition for unintended consequences that is appropriate for use in engineering design is extracted. In the fourth section, testable propositions are given, which will allow the theory to be empirically tested. Finally, the paper concludes with future directions for the theory of unintended consequences.

2 AXIOMS OF UNINTENDED CONSEQUENCES

The axioms presented in this section, which must be assumed to be true for the theory of unintended consequences to be logically consistent, are based on key features of the Law of Unintended Consequences. Specifically, the Law of Unintended Consequences implies that there is an action taken that is involved in the formation of unintended consequences. The first axiom considers whether or not unintended consequences are anticipated from the perspective of those taking this action; that is, in the context of engineering design, unintended consequences are what designers don't know or don't *want* to know.

2.1 Axiom 1: What designers don't know or don't want to know

Axiom 1. *Unintended consequences may or may not be anticipated.*

While the term *unintended consequences* is commonly attributed to Robert Merton, in actuality he used the term *unanticipated consequences* (Merton, 1936). It is worth taking a moment to consider the differences between these two terms. The word unanticipated implies that a behavior was not foreseen. In his proposed reasons why unanticipated consequences occur, Merton includes lack of knowledge, honest error, and blindness to the effect of an action on a particular group due to an ideology (cited in Dickie *et al.*, 2016; de Zwart, 2015). It is not hard to see that consequences that occur as a result of one of these reasons would be neither anticipated nor intended. On the other hand, the word unintended implies that a particular behavior was not the actor's goal or aim. For example, a medication that produces a foreseen side effect might be unintended but not unanticipated. This means that there is a class of behaviors that are both unintended and anticipated (de Zwart, 2015).

The authors propose that an unanticipated unintended consequence – what designers don't know – is fundamentally the result of the bounded rationality of designers. Bounded rational designers have limited time and mental resources and therefore may not have sufficient information to be able to make perfectly rational design decisions (Gurnani and Lewis, 2008). Simon proposed a number of reasons that bounded rational humans may not always make rational decisions (cited in Gurnani and Lewis, 2008). Of his proposed reasons, three contribute to bounded rational designers' inability to identify potential risks. These are: 1) lack of knowledge of the consequences of a decision, 2) limited ability to process information, and 3) limited time and computational resources.

For these reasons, bounded rational decision makers often use heuristics rather than optimizing design objectives directly (Lee *et al.*, 2017; Keshavarzi *et al.*, 2017). However, while heuristics are useful in practice, they often fail to create the desired system behavior. Additionally, designers may be subject to a number of cognitive biases such as those researched by Kahneman and Tversky (Kahneman and Egan, 2011; Tasic, 2009), which can further inhibit their ability to anticipate unintended consequences.

Merton suggested another source of unintended consequences that in the authors' view would make the unintended consequence anticipated – willful ignorance (Merton, 1936). Willful ignorance occurs when an action is taken even though an undesirable consequence is anticipated, if not intended. Often, the magnitude of the anticipated consequence is not known, or there may be one person in the group of decision makers who warns of a potential consequence but is ignored or dismissed by the group. Both of these scenarios would be considered anticipated unintended consequences resulting from willful ignorance.

Strictly speaking, the much more prosaic case of side effects or design trade-offs are also included under the umbrella of anticipated unintended consequences. For example, in aircraft design, there are benefits and drawbacks to choosing a T-tail configuration, including susceptibility to a dangerous deep stall condition. This drawback is anticipated, but is not willfully ignored, either. When the unintended consequence is appropriately traded-off or considered in the design, it is responsibly handled. Unfortunately, not all cases of unintended consequences are handled in this manner. By studying the mechanism by which unintended consequences form, engineers can learn to appropriately address a larger number of unintended consequences.

To summarize, according to the first axiom, the source of unintended consequences can be either 1) bounded rationality, 2) willful ignorance, or 3) known and handled appropriately. Thus, an unintended consequence can be either unanticipated, anticipated but ignored, or anticipated but managed, as shown in the conceptual diagram in Figure 1. The next axiom relates to actor(s) that actually take an action or make a decision that leads to an unintended consequence.

2.2 Axiom 2: Consequences of organized action rather than independent actions

Axiom 2. *Unintended consequences result from organized action, rather than a number of independent actions.*

The second axiom relates to the nature of the actors – specifically, that the actors are *organized* rather than having their own individual objectives. The assumption that unintended consequences result from

https://doi.org/10.1017/dsi.2019.348 Published online by Cambridge University Press

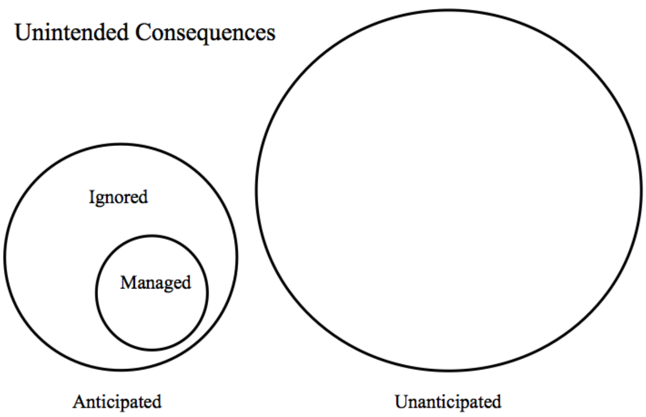

Figure 1. Unintended consequences can be either anticipated or unanticipated, and anticipated consequences can be either ignored or managed appropriately.

organized action is what distinguished Merton's work on unintended consequences from that of his predecessors and has had a large influence on the study of unintended consequences in the social sciences (de Zwart, 2015; Merton, 1936). Organized action contrasts with the actions of unorganized individuals. All actions of unorganized individuals will inherently lead to unintended consequences because individuals have their own goals unless organized (de Zwart, 2015). Engineering design has multiple organized aspects. The design process itself is organized via a set of requirements. The policy of organizations, such as Facebook's content policies, as well as regulatory oversight, such as rules limiting emissions for new automobiles, can also serve as organized action. Organized action may be a single decision such as a policy decision or a number of design decisions. While it can be difficult to trace an unintended consequence back to a single actor, groups involved in organized action – policymakers, decision makers, and designers – are assumed via Axiom 2 to be the actors whose actions have unintended consequences.

3 THEORY OF UNINTENDED CONSEQUENCES IN ENGINEERING DESIGN

In the proposed theory of unintended consequences in engineering design, the decisions of designers and policymakers are related to unintended consequences. In the following sections, unintended consequences will be defined and explanatory mathematical models will be provided.

3.1 Defining unintended consequences

Rigorously defining unintended consequences involves differentiating an unintended consequence from intended system behavior. In discussing the boundaries between an unintended consequence and intended system behavior, Merton stated the following:

> Rigorously speaking, the *consequences* of purposive action are limited to those elements in the resulting situation which are exclusively the result of the action, i.e., those elements which would not have occurred had the action not taken place. Concretely, however, the consequences result from the interplay of the action and the objective situation, the conditions of action. (Merton, 1936)

As Merton alludes, a particular behavior is not only the result of an action, but also of the nature of the system itself – the nature of the other (intended) interactions in the system as well as the state of the system. In other words, systems thinking tells us that a reductionist view in which unintended behavior is attributed to a single interaction is incomplete. It is preferable to measure the unintended system behavior (unintended consequence) as it relates to the action of the engineer, regulatory agency, or specific actor. These decision makers intend a particular system behavior when they make their decisions. Based on this assumption, the proposed measurement for unintended consequences is as follows.
An unintended consequence is the difference between intended and actual system behavior, where the intended behavior is considered from the perspective of designers, decision makers, or policymakers.
Consider the example of Facebook's role in the spread of misinformation about the Rohingya in Myanmar. So-called "fake news" about this minority group in the country was spread via Facebook and created tensions, which eventually led to real world violence against the Rohingya people (Frenkel,

https://doi.org/10.1017/dsi.2019.348 Published online by Cambridge University Press

2018). In this case, Facebook's intended behavior was that the platform would be used for connecting people and sharing information freely. The actual system behavior was this, *plus* the unintended behavior, which is the violent incidents. Therefore by the definition that has been given the violent incidents are the unintended consequence of Facebook's intentional decision not to remove misinformation.

3.2 Models of unintended consequences

Understanding that decision-making by bounded rational or willfully ignorant designers can lead to unintended consequences is not sufficient for understanding *how* unintended consequences emerge from these decisions. For explanatory purposes but also to identify design tools that could be useful for identifying unintended consequences, mathematical models of unintended consequences are necessary. The authors propose two mathematical models. Depending on the particular case, one of the two models may be more appropriate. The two models are feedback loops and bifurcations.

3.2.1 Feedback loops

Some instances of unintended consequences can be explained through feedback loops. A feedback loop is best explained via system dynamics. System dynamics is an approach based on systems theory principles that is used to model and simulate dynamic system behavior (Forrester, 1994). The power of system dynamics modeling is that it allows for the visualization of the relationships between variables as well as simulation of behavior (Fisher, 2018). A popular form of system dynamics modeling is the use of causal loop diagrams. In a causal loop diagram, system variables are connected via a directed graph. The directed edges (arrows) between variables indicate an interaction between those variables. If there is a plus sign next to the arrow, it means that when the variable at the trailing end increases, the variable at the leading end increases. If there is a minus sign next to the arrow, it means that when the variable at the trailing end increases, the variable at the leading end decreases. These interactions sometimes create feedback loops, which can be either reinforcing loops (marked with an "R" in the diagram) or balancing loops (marked with a "B" in the diagram), depending on the resulting behavior. The specifics of the interactions between variables are modeled with underlying differential equations that can be built into the model. Then, the model can be simulated. Simulation shows how the interactions between variables create a particular dynamic behavior in the system. For example, the dynamic behavior of a balancing loop is typically that a variable will reach an equilibrium point over time. The dynamic behavior of a reinforcing loop is escalation – either an upward or downward trend of a particular variable. The escalation of tensions in Myanmar due to misinformation on Facebook, discussed previously, can be explained with a reinforcing loop as in the causal loop diagram in Figure 2.

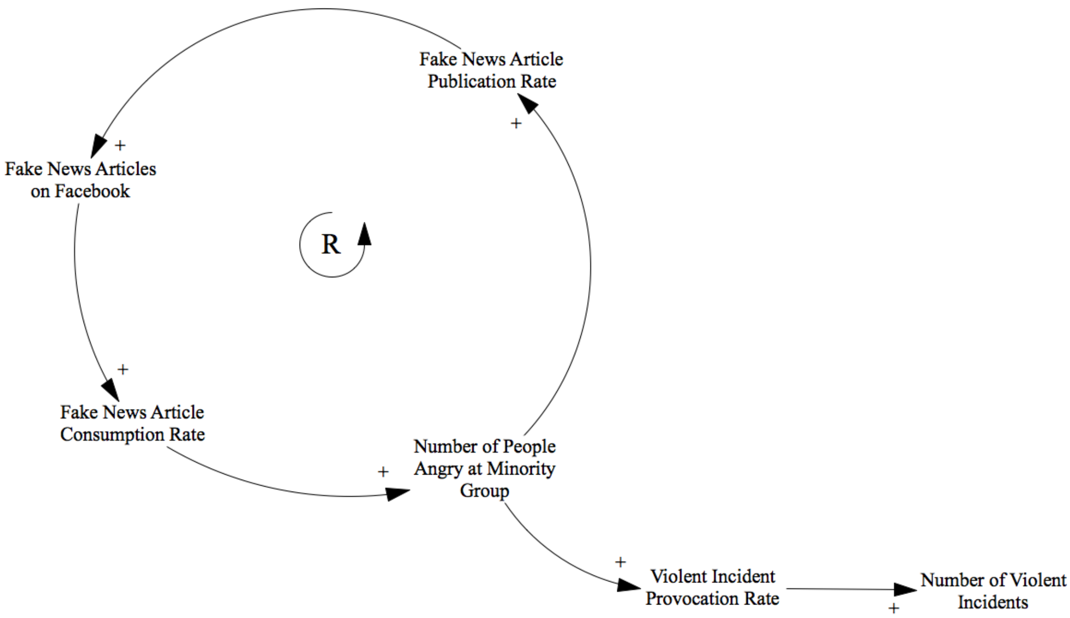

Figure 2. System dynamics modeling used to explain how facebook was exploited to incite violence in Myanmar.

https://doi.org/10.1017/dsi.2019.348 Published online by Cambridge University Press

The result of simulation of the causal loop diagram in Figure 2 is given in Figure 3. The simulation shows the change in one variable, the number of violent incidents, over time. This change in a quantity over time defines a behavior (Goel *et al.*, 2009). The unintended consequence is the escalation of the number of violent incidents over the span of time. Because the unintended consequence must only include the behavior that is the result of a decision of interest, this model accounts only for the violent incidents that were the direct result of fake news consumption. All numbers used in this example are fictitious.

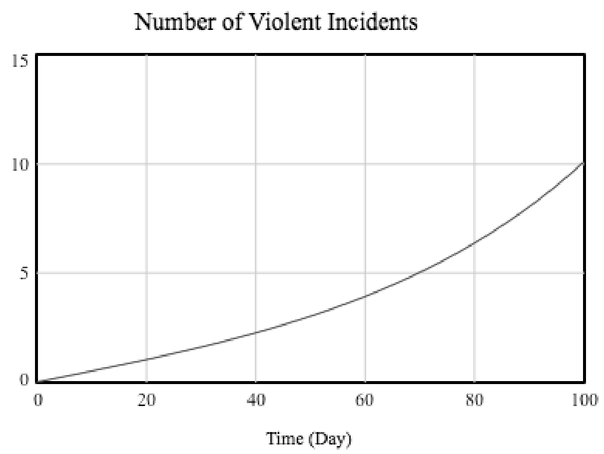

Figure 3. Simulation results from causal loop diagram from Figure 2. All numbers used in the simulation are fictitious.

Either anticipated or unanticipated unintended consequences may be described via feedback loops. Feedback may arise because of an unintended interaction between variables or because they are a part of a highly complex system in which it is difficult to predict the effects of interactions. In Facebook's case, it is likely that their leadership did not predict the degree to which misinformation spread on their platform would influence society – a product of the escalation effect shown in Figure 3 – as well as the resulting backlash against the company.

Feedback loops may explain some cases of unintended consequences. There are other cases in which an unintended consequence is not the result of a pattern such as a feedback loop, but the result of a system that operates near a critical threshold. This phenomenon can be explained through bifurcation analysis.

3.2.2 Bifurcations

Bifurcation analysis explains how relatively small actions can result in unintended consequences. Bifurcations occur when the qualitative behavior of a system changes with the value of certain parameter(s). According to Faye, a bifurcation occurs when the system suddenly changes behavior due to a small change in parameter values (Faye, 2011). For example, the stability of equilibria may change as parameter values change. Consider for example a logistic model, Equation 1, where x is a function of t and a is a constant.

$$\frac{dx}{dt} = x(a - x) \tag{1}$$

Since equilibrium occurs when $\frac{dx}{dt} = 0$, to find the equilibria of this model, the left-hand side of the equation is set equal to zero. The equilibrium points are at $x = 0$ and $x = a$. Therefore, depending on the value of a, there could be either one or two equilibrium points.

If $a = 0$, there is one equilibrium point.

If $a \neq 0$, there are two equilibrium points.

Assume a system is operating with $a = 0.1$ when a new policy changes the value of a to 0. Now, there is only one equilibrium point. This is an example of how the dynamic system's qualitative behavior can change with small quantitative changes in its parameters. The Facebook example given previously could be analyzed in a similar way using the set of differential equations underlying the causal loop diagram of Figure 2.

https://doi.org/10.1017/dsi.2019.348 Published online by Cambridge University Press

3.3 Unintended consequences: positive or negative?

In the case of Myanmar, the unintended consequence is tragic and very clearly negative. However, the ontological assumption that unintended consequences are not intended does not mean that all unintended consequences are perceived as negative. Merton asserts that unintended consequences can be perceived as either positive or negative (Merton, 1936). To illustrate this point, and to provide real-world examples of unintended consequences in a range of industries, a number of both positively and negatively perceived unintended consequences from the literature are given in Table 1, where numbers 1–9 were perceived negatively and numbers 10 and 11 were perceived positively.

Table 1. Examples of unintended consequences in the literature.

No.	Example	Reference
1	Emissions from internal combustion engine contribute to global warming	Pyper, 2014
2	Easy Bake Oven was made less efficient by regulations banning inefficient light bulbs (light bulb efficiency was measured by light production, while Easy Bake Oven used light bulbs for heat)	Schweber, 2011
3	Regulation introduced as a result of the Titanic disaster to carry more lifeboats on board decreased the stability of the ship which led to it capsizing	Stranahan, 2014
4	Creation of the internet led to the growth of cyberstalking	Williams, 2000
5	Facebook's role in the spread of misinformation led to violence in Myanmar	Frenkel, 2018
6	Tacoma Narrows Bridge collapsed as a result of an unexpected interaction between wind and structures	Larsen, 2000
7	Workplace automation can lead to carpal tunnel syndrome	Williams, 2000
8	Hospital safety feature intended to prevent medical errors caused potential for new medical errors	McDonald, 2016
9	Interaction between structures and control system on International Space Station created hazard	Gomez *et al.*, 2017
10	Adding chocolate chips to cookies, assuming the chocolate would melt and create chocolate flavored cookies, resulted in chocolate chip cookies	Gurnani and Lewis, 2008
11	Demilitarized zone between North and South Korea allowed wildlife to flourish	Kim, 1997

4 TESTABLE PROPOSITIONS

So far, it has been proposed that unintended consequences in engineering design can be modeled using system dynamics and bifurcation analysis, that they result from organized action (most commonly, design decisions or policy), and that they may or may not be anticipated. In a theoretical contribution, to specifically verify a theorized causal relationship, testable propositions are used (Whetton, 1989).

The first proposition relates to the idea of effectuation versus causation. In this context, causation is a process by which the means are chosen in order to create a certain effect and effectuation is a process by which the effect is chosen based on the means available (Sarasvathy, 2001). Sarasvathy provides a simple example to illustrate the difference. Selecting a recipe, making a shopping list, shopping, and then cooking is causation; gathering whichever ingredients are available in one's pantry and choosing a meal based on this is effectuation (Sarasvathy, 2001). The underlying logic behind effectuation is that if the future can be controlled, it need not be predicted (Sarasvathy, 2001).

Engineering design processes are generally causation-like. The effect (i.e., functions) is given via engineering requirements and design objectives, and the means are chosen accordingly. However, causation emphasizes prediction, which can sometimes lead to unintended consequences because bounded rational human designers are inherently imperfect at making predictions. To minimize unintended consequences,

https://doi.org/10.1017/dsi.2019.348 Published online by Cambridge University Press

an approach that emphasizes controlling the future rather than predicting the future is proposed to be more effective.

Proposition 1. ***Design processes that attempt to control possible effects decrease unintended consequences more effectively than processes that simply try to prevent effects.***

One reason prediction is difficult is that bounded rational designers have incomplete knowledge of a system. One situation in which knowledge is particularly limited during the design process is when designing a novel system. If it is true that lack of knowledge (one aspect of bounded rationality) is linked to unintended consequences, then it should also be true that novel designs will have more unintended consequences.

Proposition 2. ***More novel designs will have more unintended consequences than less novel designs.***

In practice, however, methods often emphasize prediction or anticipation rather than controllability. One of the main objectives of these prediction methods is to identify risks that are ignored due to group dynamics. Certain factors present in group dynamics may actually make groups less effective than individuals in certain contexts (Gyory *et al.*, 2018). Current methods for identifying risks that are unaddressed due to group dynamics include brainstorming methods that incite designers to imagine ways in which a system could fail – called a premortem (Clearfield and Tilcsik, 2018). A premortem is conducted at the beginning of a project and asks participants to consider, given that the project went wrong, why it went wrong (Klein, 2007). It has been found that people are more successful in coming up with potential risks when they imagine a hypothetical situation in which a project has failed (Clearfield and Tilcsik, 2018). In addition, project managers can also consider bringing in outsider perspectives (Clearfield and Tilcsik, 2018). Outsiders are less likely to be affected by "groupthink" and can be helpful in identifying unknown unknowns (Clearfield and Tilcsik, 2018). One particular strategy that takes advantage of this fact is called red teaming. The idea behind red teaming is to create an adversarial group whose goal is come up with ways to sabotage the project. This can be done by a real group of people, or it can be done computationally in some cases (Abbass, 2015). The benefit is that flaws in the project plan can be revealed (Abbass, 2015). Red teaming can be particularly effective in groups in which the influence of groupthink is strong or groups in which there is a very established way of doing things. These strategies can help anticipate some potential unintended consequences.

Another method that can identify potential unintended consequences is testing. Testing is important because bounded rational designers are unlikely to identify all interactions in a system, some of which may be found with testing. Testing is generally performed at different levels, from components up to system (National Aeronautics and Space Administration, 2013). At all levels, testing is performed to obtain an understanding of the couplings of the system, and the system design process should account for and, as Bloebaum and McGowan write, "embrace" these couplings (Bloebaum and McGowan, 2012). This leads to the final proposition.

Proposition 3. ***Design processes that emphasize system couplings will be more effective in reducing unintended consequences.***

Testing these propositions will contribute to the verification of the theory of unintended consequences. In addition to empirical verification, there is much work that remains to be done in developing design tools for the identification of unintended consequences. One of the proposed methods comes from one of the models that was used in this paper to explain how unintended consequences form: system dynamics.

5 TOWARDS THE IDENTIFICATION OF ARCHETYPES FOR UNINTENDED CONSEQUENCES

There is a useful tool in system dynamics for recognizing patterns of interactions between variables that lead to a particular behavior. System archetypes are essentially motifs which have been observed in multiple cases and which lead to a particular behavior (Marais *et al.*, 2006). As an example, the commonly-used tragedy of the commons scenario is a system dynamics archetype. A tragedy of the commons occurs when individual actors each take an action that is best for their individual needs, but the collective effect of a large number of actors making these decisions sabotages those individual effects for all actors. One example is antibiotic resistance. The recognition of a particular archetype allows decision makers to put policies in place or re-architect the system in order to mitigate or avoid the effect of the archetype.

https://doi.org/10.1017/dsi.2019.348 Published online by Cambridge University Press

Future research will identify system archetypes that lead to unintended consequences. It is also possible to suggest strategies for re-architecting the system based on the particular archetype. Wolstenholme suggests that for every archetype, there is a specific archetypal solution (Wolstenholme, 2003). It would be a useful design tool to have a reference list of archetypes that contribute to unintended consequences as well as a suggested solution archetype for each. This approach, combined with novel design methodologies that emphasize controllability rather than predictability, will be the focus of future research.

Months after a number of scandals related to Facebook came to light, including what happened in Myanmar, it came out that Facebook had known about some of its issues earlier than it had previously admitted but ignored warning signs and attempted cover-ups (Frenkel *et al.*, 2018). Given that novel technologies often have effects that are difficult to foresee, it was not altogether surprising that there would be at least one unintended consequence of social media. However, as the saying goes, the cover-up is often worse than the crime, and Facebook's actions after uncovering unflattering consequences of its policies were no exception. The Law of Unintended Consequences tells us that unintended consequences are inevitable; the question is, what can we do to manage them responsibly?

REFERENCES

Abbass, H. (2015), *Computational Red Teaming*, "Springer International Publishing". http://dx.doi.org/10.1007/978-3-319-08281-3

Bloebaum, C. and McGowan, A. (2012), "The design of large-scale complex engineered systems: present challenges and future promise", in *Proceedings of the 12th AIAA/ISSMO multidisciplinary analysis and optimization conference*, Indianapolis, IN, USA, 17–19 September, AIAA. http://dx.doi.org/10.2514/6.2012-5571

Clearfield, C. and Tilcsik, A. (2018), *Meltdown: Why our systems fail and what we can do about it*, Penguin Publishing Group.

Davenport, C. (2018), "Washington Rolls Back Safety Rules Inspired by Deepwater Horizon Disaster", *The New York Times*, 27 September.

de Zwart, F. (2015), "Unintended but not unanticipated consequences", *Theory and Society*, Vol. 44 No. 3, pp. 283–297. http://dx.doi.org/10.1007/s11186-015-9247-6

Dickie, M., Gozdecka, D. and Reich, S. eds. (2016), *Unintended Consequences: The impact of migration law and policy*, ANU Press. https://doi.org/10.22459%2Fuc.08.2016

Faye, G. (2011), *An introduction to bifurcation theory*, NeuroMathComp Laboratory, Paris, France.

Fisher, D. (2018), "Reflections on teaching system dynamics modeling to secondary school students for over 20 years", *Systems*, Vol. 6 No. 12. https://doi.org/10.3390%2Fsystems6020012

Forrester, J. (1994), "System dynamics, systems thinking, and soft OR", *System Dynamics Review*, Vol. 10 No. 2–3, pp. 245–256. https://doi.org/10.1002%2Fsdr.4260100211

Frenkel, S. (2018), "Facebook to Remove Misinformation That Leads to Violence", *The New York Times*, 18 July.

Frenkel, S., Confessore, N., Kang, C., Rosenberg, M., and Nicas, J. (2018), "Delay, Deny and Deflect: How Facebook's Leaders Fought Through Crisis", *The New York Times*, 14 November.

Goel, A., Rugaber, S., and Vattam, S. (2009), "Structure, behavior, and function of complex systems: The structure, behavior, and function modeling language", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 23 No. 1, pp. 23–35. http://dx.doi.org/10.1017/s0890060409000080

Gomez, S., Platonov, V., Medina, E., Borisenko, A., and Bogachev, A. (2017), "Unexpected Control Structure Interaction on International Space Station", in *International Astronautical Congress 2017 (IAC 2017)*, 25-29 Sepember, Adelaide, Australia.

Gurnani, A., and Lewis, K. (2008), "Collaborative, Decentralized Engineering Design at the Edge of Rationality", *Journal of Mechanical Design*, Vol. 130 No. 12, p. 121101. http://dx.doi.org/10.1115/1.2988479

Gyory, J., Cagan, J., and Kotovsky, K. (2018), "Should teams collaborate during conceptual engineering design?: an experimental study", in *Proceedings of the ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Quebec City, Canada, 26–29 August, ASME. http://dx.doi.org/10.1115/detc2018-85602

Kahneman, D. and Egan, P. (2011), *Thinking, fast and slow*, Farrar, Straus and Giroux, New York, NY, USA,

Kim, K. (1997), "Preserving Biodiversity in Korea's Demilitarized Zone", *Science*, Vol. 278 No. 5336, pp. 242–243. https://doi.org/10.1126/science.278.5336.242

Kim, S. (2012), "Characterizing unknown unknowns", in *PMI Global Congress 2012–North America*, Vancouver, BC, Canada, PMI, Newton Square, PA, USA.

Keshavarzi, E., McIntire, M., Goebel, K., Tumer, I., and Hoyle, C. (2017), "Resilient System Design Using Cost-Risk Analysis With Functional Models", in *Proceedings of the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Cleveland, OH, USA, 6–9 August, ASME. http://dx.doi.org/10.1115/detc2017-67952
Klein, G. (2007), "Performing a project premortem", *Harvard Business Review*, Vol. 85 No. 9, pp. 18–19. http://dx.doi.org/10.1109/emr.2008.4534313
Larsen, A. (2000), "Aerodynamics of the Tacoma Narrows Bridge – 60 years later", *Structural Engineering International*, Vol. 10 No. 4, pp. 243–248. http://dx.doi.org/10.2749/101686600780481356
Lee, B., Fillingim, K., Binder, W., Fu, K., and Paredis, C. (2017), "Design Heuristics: A Conceptual Framework and Preliminary Method for Extraction", in *Proceedings of the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Cleveland, Ohio, 6–9 August, ASME. http://dx.doi.org/10.1115/detc2017-67467
Marais, K., Saleh, J., and Leveson, N. (2006), "Archetypes for organizational safety", *Safety Science*, Vol. 44 No. 7, pp. 565–582. http://dx.doi.org/10.1016/j.ssci.2005.12.004
McDonald, C. (2006), "Computerization Can Create Safety Hazards: A Bar-Coding Near Miss", *Annals of Internal Medicine*, Vol. 144 No. 7, pp. 510–516. http://dx.doi.org/10.7326/0003-4819-144-7-200604040-00010
Merton, R. (1936), "The Unanticipated Consequences of Purposive Social Action", *American Sociological Review*, Vol. 1 No. 6, pp.894–904. http://dx.doi.org/10.2307/2084615
Mullins, J. (2010), "The eight failures that caused the Gulf oil spill", *New Scientist*, 8 September.
National Aeronautics and Space Administration (2013), *Systems Development, Integration & Test* [online]. Available at https://www.nasa.gov/sites/default/files/atoms/files/g_28367u_sdit_es60.pdf (Accessed: 1 November 2018)
Pyper, J. (2014), "Car, Truck and Airplane Pollution Set to Drive Climate Change", *Scientific American*, 10 April.
Ring, J. and Madni, A. (2005), "Key challenges and opportunities in 'system of systems' engineering", in *2005 IEEE International Conference on Systems, Man and Cybernetics*, Waikoloa, HI, USA, 12 October, IEEE. http://dx.doi.org/10.1109/icsmc.2005.1571272
Rocha, V., Ries, B., and Wills, A. (2018), "Mark Zuckerberg testifies before Congress", *CNN Politics*, 11 April.
Sarasvathy, S. (2001), "Causation and Effectuation: Toward a Theoretical Shift from Economic Inevitability to Entrepreneurial Contingency", *Academy of Management Review*, Vol 26 No. 2, pp. 243–263. http://dx.doi.org/10.2307/259121
Schweber, B. (2011), "The law of unintended consequences strikes, again", *EE Times*, 22 September.
Stranahan, S. (2014), *The Eastland Disaster Killed More Passengers Than the Titanic and the Lusitania. Why Has It Been Forgotten?* [online]. Available at: https://www.smithsonianmag.com/history/eastland-disaster-killed-more-passengers-titanic-and-lusitania-why-has-it-been-forgotten-180953146/ (Accessed: 26 October 2018)
Tasic, S. (2009), "The illusion of regulatory competence", *Critical Review*, Vol. 21 No. 4. PP. 423–436. http://dx.doi.org/10.1080/08913810903441369
Whetton, D. (1989), "What constitutes a theoretical contribution?", *Academy of Management Review*, Vol. 14 No. 4, pp. 490–495. http://dx.doi.org/10.5465/amr.1989.4308371
Williams, J. (2000), "Systems thinking and unintended consequences", *ADTMag*, 21 August.
Wolstenholme, E. (2003), "Towards the definition and use of a core set of archetypal structures in system dynamics", *System Dynamics Review*, Vol. 19 No. 1, pp. 7–26. http://dx.doi.org/10.1002/sdr.259

ACKNOWLEDGMENTS

This research is supported by the National Science Foundation (NSF CMMI 1562027). Any opinions or findings of this work are the responsibility of the authors and do not necessarily reflect the views of the sponsors or collaborators.