

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

Better anticipating unintended consequences

Clinton J. Andrews, *Life Member, IEEE*

Abstract— If people want the benefits of innovations, must they simply accept the unintended adverse consequences? Versions of this question haunt many who care about the social implications of technology. Technological design processes could include impact assessment steps, but not all do. Adoption in the marketplace may ignore spillover effects. Jurisprudence is often reactive and focused on remediating obvious wrongs. Public policy also often requires evidence of harm before legislators or administrators are willing to act. The failure to anticipate adverse consequences is sometimes framed as a moral lapse, but it could equally be about competence or incentives. This paper considers the relative merits of methodology (analogizing, interpolating, projecting,) and procedure (reflecting, reasoning, discourse) as systematic approaches to anticipating unintended consequences of innovation. It weighs the efficacy of such approaches against current reactive remedies, highlighting the importance of tailoring approach to context, and building in early learning opportunities (observing and testing). Several examples suggest that society is often playing catch-up and trying to avoid adverse consequences before the innovation is widely deployed rather than before it is initially introduced.

Index Terms—Futurism, philosophical considerations, risk analysis, social implications of technology, technology forecasting, uncertainty, unintended consequences

I. INTRODUCTION

A surprise is an unexpected event or thing. The nature of surprises has changed since Shakespeare [1] embodied them in his mischievous character Puck. Adverse historical surprises were mostly natural disasters, plagues, wars, and plots, whereas today many are technological. We invent new things, deploy them at scale, apply them in new ways, and discover surprising outcomes, some adverse.

Unexpected and unintended outcomes are distinct effects. Unexpected benefits, unexpected drawbacks, and perverse results of purposeful social action are all unintended [2]. However, not all unintended consequences are unanticipated. For example, when dispensing a prescription for a pain-reducing drug that also induces drowsiness, a doctor will advise the patient to avoid driving to reduce the odds of an unintended but expected adverse outcome. Approaches for reducing unintended consequences therefore include efforts to reduce surprise or the unexpected, and to reduce the odds of adverse consequences.

This paper argues that the ability to manage unintended adverse consequences of technological innovations depends on discussing them more thoroughly. However, some modes of discussion are more effective than others. Effective discussion has both substantive and procedural dimensions that this paper unpacks. The substantive challenge is to contribute insights based on current knowledge to inform scenarios for a future in

which the technology is deployed at scale. The procedural challenge is to get the right people discussing the technology and its potential consequences at times and places where they can exert leverage on the technology's trajectory. This paper draws on a literature in communicative rationality, technology policy, and interpretation of illustrative examples to suggest what the desired "more thorough discussion" might mean in practice.

This paper first discusses its methods, and then the institutions societies use to help respond to adverse consequences, and the risk-benefit tradeoff calculus underlying efforts to prevent such consequences. The institutional description focuses on the U.S. context to illustrate more general considerations, while acknowledging that national frameworks vary. The paper then examines the substantive and procedural dimensions of discussions seeking to anticipate unintended adverse consequences of technological innovations, considering the promise of specific strategies. It highlights the value of quickly learning from experience early in the technology deployment process. The paper applies its communicative analysis framework in an illustrative manner to four widely different historical examples including chlorofluorocarbons, drones, Thalidomide, and Facebook's Newsfeed. It draws tentative conclusions regarding the contributions of specific substantive and procedural strategies and recommends next steps.

II. METHODS

This paper employs qualitative research methods that focus on how people interpret interactions among evolving science, technology, and policy narratives. Following [3], it investigates public policy development as a process of developing a narrative: "policy narratives are stories (scenarios and arguments) which underwrite and stabilize the assumptions for policymaking in situations that persist with many unknowns, a high degree of interdependence, and little, if any, agreement." This methodology involves reading and interpreting texts to extract their meanings (and not hubristic and implausible social 'laws') [4], and which "privilege meanings as ways to grasp actions" [5] while acknowledging the contingency of such context-specific insights. The paper employs a conceptual framework originating in theories of communicative action [6] and critical policy analysis [7]. The short case examples interpret the historical literature to extract suggestive rather than conclusive insights.

III. RESPONDING TO ADVERSE CONSEQUENCES

Social theorists such as Marx, Weber, Merton and Sartre have speculated that most human interactions yield unintended consequences [8]. Many have catalogued unintended adverse

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

outcomes from the accumulated impacts of many small events, the system features that emerge from individual interactions, changes in contextual factors, spillover effects such as network externalities, and dual-use applications in fields of innovation such as energy, environment, medicine, and urban planning, [e.g., 9-12]. Society's institutions regularly respond.

Humans respond to problems both individually in personal decisions, and collectively through participation in systems of economic and political decision making.

As individuals, humans operate in a world that they do not fully understand by relying on habits, heuristics, and biases that work sometimes but not all the time [13]. Their mental models are incomplete and inconsistent [14]. Neuroscientists studying individual humans identify “puzzlement” surprise, measured by how unexpected an event is, that may trigger changes to habitual (model-free) behavior. This differs from “enlightenment” surprise, measured by “how much an event changes our model of the world and, as a consequence, our expectations,” that may trigger changes to reasoned behavior and encourage pursuit of better information [15]. People carry their individual capacities and biases to social roles as consumers and citizens.

Markets are relatively efficient, voluntary and adaptive mechanisms for allocating societal resources, and many innovations diffuse through marketplace choices. Innovations that exhibit relative advantage, compatibility with existing assets, low complexity, trialability, and observability are likely to diffuse quickly [16]. Each of these characteristics represents a means by which a potential adopter can assess benefits and costs of the innovation. Many bad ideas fail the market test.

Political institutions set the rules for marketplace and societal actors. Most modern governments have a constitutional basis that assigns roles to each branch. For example, in the U.S. context, legislatures promulgate laws, the executive branch implements laws, and the courts adjudicate the constitutionality of laws as implemented. When innovations that succeed in the marketplace turn out to produce negative spillovers or unintended adverse consequences, aggrieved parties can seek governmental redress.

Often, a first point of entry is the court system, where one party can seek relief from harm caused by another party. A successful lawsuit under tort liability (a widely used construct with Roman origins) typically requires evidence of harm, hence it is a reactive mode of redress that attempts to cure a problem after the fact [17]. In U.S. parlance, the most relevant torts for the purposes of this paper are negligence (breach of a duty of care) and nuisance (activity that harms others).

The unhappy reality is that legislative and executive branch solutions are also often reactive. Issues usually must climb systemic agendas (by garnering newspaper headlines, airtime, and organized advocates) before ever getting onto institutional agendas (legislative hearings, legislation, executive actions) [18]. Modern social media technologies simultaneously raise the possibility of more democratic, bottom-up [19] and more top-down, algorithmic forms of agenda-building [20].

Designers' good judgement, organizational concerns about reputation and liability, market forces and governmental interventions have reduced some of the adverse consequences experienced by society from technological innovation. But new problems continually emerge. Society's institutions therefore scan for and assess emerging risks.

Risk assessment is the applied science of anticipating the likelihood of adverse consequences, and risk management is the practice of reducing them to the extent that society supports [21]. The utilitarian, consequence-focused framing of this practice has been criticized for hubris (its claims of objectivity, precision and comprehensiveness) and for embedding unreasonable decision rules (such as net social benefit, with winners and losers) in rational methodology [22, 23]. Nonetheless, risk-based policies are widely adopted by governments and enterprises. The challenge has been to build the political will to adopt timely policies.

Precautionary approaches change the basis for policymaking to emphasize risk avoidance over risk management. They have been widely adopted in Europe but not North America. Precaution depends on similar risk assessment analytics for anticipating adverse consequences, but uses the knowledge gained to determine whether to allow an innovation to deploy rather than whether to regulate the extent and circumstances of that deployment [24]. Some critics see this as a distinction without a difference because the assessments and management activities are so similar [25]. Others are concerned that excessive precaution (often by government) stifles innovation (by market actors with an appetite for risk) and hence call for innovation systems that “counterbalance the generation of novelty and the risk of negative unintended consequences” [26].

The human ability to anticipate adverse consequences remains limited enough that both precaution and risk management are necessary in practice. The Amish are well known for deliberating in a precautionary way before allowing members of their community to adopt new technologies, yet they also manage other risks—such as weather-related crop failure—in a very consequentialist way [27]. At a larger scale, foresight-oriented activities such as technology roadmaps lay out an organization's or sector's technology adoption strategy, and deliver insights that benefit both risk management and precautionary management approaches. Research and development investments in particular capture the goal of being anticipatory rather than reactive or precautionary.

The standard tools for managing unintended consequences thus are largely reactive. Such tools are practical because they can be evidence-based and are less likely to waste society's resources by responding to insignificant concerns. However, they are morally and intellectually unsatisfying. Is it possible to do better by attempting to anticipate adverse effects?

A search for better ways to anticipate unintended adverse consequences is also a quest for more rational thinking about the future. This paper follows [28] in distinguishing between substantive and procedural dimensions of rationality. “Behavior is substantively rational when it is appropriate to the achievement of given goals within the limits imposed by given conditions and constraints” [p. 66] and “behavior is procedurally rational when it is the outcome of appropriate

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

deliberation” [p. 67]. Because deliberation often involves multiple parties, many call for a communicative rationality that seeks for every deliberation a social context in which participants can debate the “truth, rightness, and sincerity” of claims [29]. These concepts have entered practice in the form of varied creative efforts to marry substance and process including analytic deliberation [30], participatory technology assessment [31], and joint fact-finding [23].

IV. SUBSTANTIVE APPROACHES FOR ANTICIPATING UNINTENDED ADVERSE CONSEQUENCES

Predicting the future is something people do regularly, but they are better at predicting some events than others. After decades of investment in data collection networks and modeling systems, people now have become pretty good at predicting tomorrow’s weather, but not next month’s weather. Although people have invested vast effort in data and models to predict tomorrow’s stock market prices, they often miss wildly.

Uncertainty bedevils predictions. Relevant types include statistical variation, subjective judgment, linguistic imprecision, variability, inherent randomness, disagreement, and approximation [30]. Some uncertainties are irreducible, such as the inherent unknowability of the future or the actions other entities might take in response to current conditions, but people can address some types of uncertainty with improved data, tools, relationships, and wisdom.

In situations where uncertainties are rife and predictions are often inaccurate, the focus shifts toward management of the associated impacts. Anticipating surprise is a fundamental survival skill requiring adaptability and foresight. Adaptability, the quality of being able to adjust to new conditions, improves with access to diverse resources, internal capacity, and a flexible mindset [33, 34]. Foresight, the ability to consider the future in a systematic way, improves when actors identify salient issues or concerns, track their development over time, and anticipate changes in their trajectories [35]. Futurists warn that “predictable surprises” happen when we fail to recognize salient factors, track trends, and commit resources in time [36]. Accurate assessments of what matters (saliency) are difficult because people may fail to focus on important goals due to the many distractions they experience [37]. Additionally, they may encounter **un**predictable surprises.

Prediction seeks to identify a specific, most likely future, and forecasting attaches a date when that future will occur. Foresight instead seeks to help people understand the ramifications of several plausible futures so that they can manage the consequences [38].

Anticipating consequences requires a logical basis for projecting what might happen, given what has already happened and what people know about how the world works. It requires argumentation about cause and effect, and it requires a system conception capable of delivering a “then” when given an “if.” A simple physical system may have more data available and more tractable modeling tools to characterize it than would a complex sociotechnical system. Thus, system representations vary in formality and usefulness for both prediction and foresight.

Disciplines integrate causality, probability, and mechanism in different ways that make multidisciplinary communication both

valuable and challenging [39]. Futurists distinguish among anticipated-intended consequences, anticipated-unintended consequences, and unanticipated-unintended consequences, and their methods seek to turn unanticipated-unintended into anticipated-unintended consequences [12]. Such methods identify relevant events, trends, and disruptions using means that vary from informal to formal, depending on data availability, analytical resources, time horizon, technological status, and other factors.

It is useful to array methods for characterizing systems on a spectrum, from analogizing, to interpolation, to projection. Each method has strengths and weaknesses, and each exhibits better fitness for some applications than for others, as discussed below and summarized in Table 1. These methods illustrate but do not exhaustively cover the spectrum of possibilities.

A. Analogize

Are there relevant historical analogies from which we can learn? What are the limits of analogizing? Examples include: (1) “the old network was like a narrow country road, but the new one will be like a superhighway” [40]; and (2) the “paperworld” permitted by the commodification of books due to invention of the printing press presages problems in today’s cyberspace [44].

Writers distinguish rhetorical gradations from simile (this is like that), to analogue (this is a stand-in for that), to metaphor (this is that). Metaphor is usually poetic, but analogy is often literal and quite useful for thinking about ‘what-if’ scenarios. Speculative fiction relies heavily on analogy [41].

Analogical reasoning is a cognitive heuristic—a shortcut—that is widely used to make sense of the world [42]. Perceived similarities between cases become the basis for claiming that further similarities may exist. For example, scientists regularly perform laboratory experiments on mice to predict likely human responses to biomedical interventions. Good biomedical researchers also understand the limits to analogy and thus conduct follow-up studies with human subjects.

Logicians note that the strength of analogical reasoning depends more on the content of the argument than on its form, that is, the extent to which similarities between cases are present and relevant, and not offset by inevitable dissimilarities [43]. If the measure of an analogy is its empirical strength, then one necessarily needs to evaluate it in a specific context. Historians and other humanists specialize in contextualized interpretation when crafting and contesting analogies [44].

There is an ethical dimension, because while “analogical imagination” has great value in exploring possible futures for emerging technologies, users may wield specific analogies rhetorically to advance strategic arguments, leading to the need to develop a critical “analogical sensibility” [45]. The practical implications are to appreciate similarities, dissimilarities, and limits to generalizability when deploying any analogy, and to encourage stakeholders to offer multiple analogies as they deliberate.

B. Interpolate

How does a fine-grained appreciation of the dynamics of social practices guide us to unbundle social meanings and

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

human competencies from old technologies and successfully re-bundle them with new technologies? An example is: “Patient mortality increased significantly...following introduction of a new computerized physician order entry system aimed at reducing medical error...[because the system] did not perform as anticipated in real-world situations” [46: pp. 1-2].

Social and behavioral considerations at the level of individual technology adoption decisions can scale up to be just as important as those affecting systemwide technology diffusion. Anthropologists and sociologists have documented strong determining roles for everyday social practices that cause naturalistic adoption behaviors to differ from those observed in controlled experimental settings [47]. One useful framework defines a social practice as bundle incorporating a material element (a technology), a procedural element (competencies and social rules needed to use the technology), and a meaning element (symbols, ideas, and aspirations people attach to the technology) [48]. Changing a social practice involves unbundling all three elements and then re-bundling a new set of material, procedural, and meaning elements into a new social practice. Thus, an apparently simple technology swap, say, adopting an electric car in place of one with an internal combustion engine, is also likely to require the user to develop new procedures (such as learning to use regenerative braking and managing limited driving range) and meanings (such as re-imaging the prestige of a luxury brand or the ruggedness of a pickup truck with an 8-cylinder engine).

A common refrain from social practice researchers is to expect new technology adoptions to progress more slowly than a simple technology substitution analysis might predict, because people only slowly develop new procedures and assign new meanings. Technology developers therefore ought to ask how close a substitute to what it replaces is the new technology, how user friendly is it, and how do its social meanings diverge from the status quo? If technology adoption is in fact a highly social process, then it makes sense to use social knowledge and tools, from consumer education to peer pressure, to encourage adoption. This is also an opportunity to identify unintended consequences.

C. Project or Extrapolate

How do our understandings of physical, chemical, biological, economic, social, and political system dynamics help us predict the systemic effects of widespread deployment of an innovation? An example is: “We estimate robust negative effects of robots on employment and wages” [49: p. 1].

Researchers have long been interested in the dynamics of natural-science, economic, social, and political systems, and in recent decades the toolkit has improved dramatically. The systems approach was embraced by military technology planners in the 1950s and 1960s, reviled due to misconceived efforts to apply these tools to civil problems in the 1970s and 1980s, and rehabilitated as computing and visualization technologies opened new possibilities in the 1990s and 2000s [50]. In economics, anecdotes and extrapolations from historical narratives to support conceptual models (such as Marx’s political economy) have been joined by mathematical modeling of simplified systems (such as Walras’ general equilibrium model of the macroeconomy), quantitative

documentation of observed economic system dynamics (such as Schumpeter’s business cycles and Kondratiev’s long waves) and structures (such as Leontief’s sectoral demand model), and formal modeling of dis-equilibrium conditions (such as Schelling’s agent-based models). High interest from governments and businesses, and an abundance of detailed time series data have fueled these advances in economic modeling and projections.

Social and political system dynamics are less well understood and there is less of a tradition of predictive modeling. This is due to a lack of consensus on appropriate conceptual models of these systems, a lack of adequate time series data, the contingent nature of social knowledge, and the high importance of contextual factors [7, 51].

Integrated modeling to predict the impacts of technological innovations on society nonetheless has deep roots. The intellectual framework of public choice theory advocated by Black, Arrow, Downs, Olsen, Buchanan, Tulloch, and others portrays many social, political, and economic phenomena as the outcome of interactions among individual, essentially microeconomic actors [52]. Contrasting theories that attribute causality to structural factors in society, whether Marx’s class system or Carmichael’s institutional racism, can be folded into an integrated assessment using the logic of structuration [53]. Structuration theory asserts that agency and structure are interdependent because structure constrains agency, even as agents must actively decide to reproduce structure. A multi-level perspective has gained currency in recent years to acknowledge multiple levels of resolution in both agency and structure [54]. Agent-based modeling has been useful for formalizing these concepts and exploring observed implications such as tipping points [55, 57].

Integrated assessment is widely used for public policy analysis, and it is most advanced in applications to climate change policy [57]. These models marry bottom-up technology characterizations, such as types of electric power generation plants, with technology substitution models that predict their market share over time, with top-down macroeconomic models that account for changing prices and quantities, as mediated by policies and fundamental resource constraints, and the resulting system simulations report salient metrics such as net carbon emissions and changes in Gross Domestic Product. The virtue of an integrated assessment is its internal consistency and comprehensive scope. An unwanted byproduct is the extreme level of uncertainty associated with any prediction because so many factors are involved. Thus, integrated assessments are most useful for “what-if” scenario analysis to explore a decision space.

Many technological innovators lack the resources to conduct formal, integrated modeling of social, economic, and political system dynamics as part of the design process. Instead, they use approximations and rules of thumb. For example, if introducing a more efficient technology reduces the price to consume a unit of output, then consumers might use the money saved to consume even more of the product, causing a rebound effect that takes back some of the net resource savings [58]. If the efficiency improvement is dramatic enough, when it scales up it can change the entire economic structure, as was the case with 18th-century steam engines and 20th-century integrated circuits.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

Relevant metrics in these cases are the price and the price sensitivity of demand for the product. Each rule of thumb analyzes only a small part of the integrated system to assess impact.

Whether the sociotechnical system model is integrated or piecemeal, open or closed, explicit or reduced form, it can only rarely claim to be comprehensive across all relevant phenomena. Exogenous factors affect nearly all system models. Thus, there are inevitable limits to the value of these models in projecting future outcomes. Communicating these limits to users and asking whether the model is fit for its purpose can improve the odds that it makes a substantive contribution to the investigation of unintended consequences.

V. ADDING A PROCEDURAL DIMENSION

Systematic thinking is only part of what society needs to anticipate adverse consequences of technological innovations. There is also a procedural dimension that assigns responsibilities to certain people and organizations. One can distinguish between “technologies of hubris” (predictive methods) and complementary “technologies of humility” that identify unforeseen consequences, admit normative biases, and promote pluralistic discussion [59]. Such a communicative rationality requires a marriage of substance and process. Lively literatures on anticipatory governance [e.g., 60] and responsible innovation [e.g., 61] attempt to establish procedural norms within innovation ecosystems [62]. Procedures comprise a spectrum spanning individual reflection, moral reasoning, and discussion. Each has strengths, weaknesses, and appropriate applications, as discussed below, and summarized in Table 1. Again, the list is not exhaustive.

A. Reflect

What can we learn from engaging in reflective professional practice? How can we remember to look back and assess our own work periodically? An example is: “The [nuclear] accident at Three Mile Island was caused by a combination of hard-to-handle machinery, poorly trained or incompetent operators, and a regulatory process that lulled management into neglecting its own responsibilities. Congress, the public, the regulators, and the nuclear power industry reacted. Some overreacted...” [63]

The everyday professional activities of engineers and designers are social practices with high relevance to new technology development. By reflecting on designers’ practical actions, learning may occur, and outcomes may improve [64]. Observers note that experience alone does not guarantee better performance, because it also requires deliberate reflection on the experience [65]. The act of reflecting can be unsettling because it raises the possibility that one has some personal responsibility for what one’s designs do to the world, that the social dimensions of professional work—from employee diversity to managerial style—might affect outcomes, and that living with uncertainty and doubt are intrinsic features of professional practice [66].

Design work involves making tradeoffs, and each tradeoff decision brings the potential for unintended consequences. Reflective practice provides opportunities to revisit those tradeoff decisions and play out what-if scenarios. Whether institutionalized as an “after-action debriefing” or undertaken

as a personal journey, reflection enhances understanding. It pursues answers to three inter-related questions: “What?” “So what?” and “Now what?” that detect unintended consequences [67].

B. Reason

Are there usable frameworks for moral reasoning about professional practice? An example is: Principles that should govern age-appropriate digital services include “presenting information in an age-appropriate way, upholding children’s rights, offering fair terms for children, recognizing childhood, and putting children ahead of commercial interests and ahead of platform status” [68].

Moral reasoning by individual professionals can help them anticipate unintended consequences. The personal guidance of an ethical framework, whether motivated by concern over potential consequences or a sense of mutual obligation, can encourage timely questioning of motives and consideration of outcomes. This personal ethical stance might be guided unconsciously by deeply held assumptions, informally by social norms, or formally by regulative constraints.

Culture, defined as the shared attitudes, beliefs, practices, and values that characterize a group [69], instills within each person a set of unwritten framing assumptions that influence their behavior. Each person also has cognitive biases that limit their ability to think logically, comprehensively, and quickly.

Social norms are patterns of behavior that are self-reinforcing within a group and may infer an evolutionary advantage on members of the group [70]. Informal but explicit social norms may powerfully constrain individual behavior. Many professions and organizations write ethical codes to guide member behavior. Some elements merely protect the group from ethical lapses by participating individuals, but most also advance some vision of “the public good.” When the consequences of behavioral lapses have the potential to become significant, informal normative constraints give way to formal legal and policy constraints [71].

Moral reasoning involves asking good questions. Thus, one professional code of ethics asks members how they “hold paramount the safety, health, and welfare of the public” and “improve the understanding by individuals and society of the capabilities and societal implications of conventional and emerging technologies” [72]. Such questions can help a designer to weigh tradeoffs and reduce adverse consequences. This specificity might be as or even more useful in a practical context than the overarching injunction to avoid doing harm [73].

C. Discourse

Can talking about planned new technologies help designers avoid adverse outcomes? An example is: “Deliberation enabled the participants to develop a shared understanding of New Jersey’s energy future” [and] “forged a consensus on targets for three key energy indicators” (greenhouse gas emissions, percent of income spent on energy, local self-sufficiency) while recognizing “that the future will be created by active and concerned citizens, consumers, businesses, workers, and investors.” [74: pp. 18-19]

Opening an inquiry up to broader participation can introduce more varied points of view and bring more types of knowledge

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

to bear on a problem, provided participants communicate successfully [75]. Communication can fail when one or more participants decline to participate voluntarily in the discussion, or when distortions hinder effective communication [76]. In complex technical domains, just talking might not deliver enough rigor, hence it may become necessary to marry the instrumental (technical analysis) and communicative (open discussion) dimensions of rationality by using careful processes of co-design [77], joint fact-finding [23], or analytic deliberation [74].

Open discussion of the social implications of new technologies is such a well-established value in social studies of science and technology that people sometimes forget how difficult it is to do well [78]. Each emerging technology presents specific challenges to effective discourse. For example, the opacity of current AI has limited “deep” discussion of its merits and demerits [79]. Yet the desire to communicate, the norms of ethical inquiry, and the value of discourse persist even as the deliberations take new forms and move online [80].

VI. LEARNING BY DOING

There are limits to the human ability to anticipate unintended outcomes in the real world. An important role for learning quickly remains. The slow, reactive form of social learning embedded in governance processes such as lawsuits, legislation, and regulation may not be timely. Observation and testing are tools that speed up the learning process.

A. Observe

If adverse consequences are suspected, can they be detected promptly even as the innovation deploys at scale? An example is: “Although the need for early drug safety monitoring is widely recognized, a harmonized approach at a global level is still lacking” [despite] “the globalization of the clinical landscape over recent decades” [81].

Science advances through an iterative process of observing phenomena, making empirical generalizations, formulating theories, proposing testable hypotheses, and carrying out those tests through observation [82]. Repeated, time-series observations are valuable for detecting change in dynamic systems and allowing the “wheel of science” to spin rapidly and advance knowledge. Thus, if adverse consequences of deploying an innovation are suspected, then it makes sense to set up a surveillance process capable of detecting such consequences.

Surveillance processes are longstanding in the fields of public health and environmental management. One of the most challenging aspects is to figure out what to observe. In public health, it is customary to do surveillance planning to identify hazards warranting attention [83] and in environmental management it is common to screen candidate contaminants of emerging concern using existing knowledge of hazard characteristics and exposure potential [84]. Both fields receive regular governmental support for surveillance activities although they are chronically under-funded relative to their scopes of concern [8].

B. Test

Is it possible to perform low-risk tests to identify potential adverse consequences of innovations before they are deployed at scale? An example is: “Online education is rapidly expanding in response to rising demand for higher and continuing education, but many online students struggle to achieve their educational goals...Scaling behavioral science interventions across various online learning contexts can reduce their average effectiveness by an order-of-magnitude. However, iterative scientific investigations can uncover what works where for whom.” [86]

Designers typically test many alternatives before settling on a preferred solution, so experimentation is already a core element of innovation. Unlike the role of experimentation in science, the focus of a design experiment is “not to definitely accept or reject a hypothesis, but rather to iteratively refine the intervention” [87: pp. 163–164]. Many of these tests will be low risk, performed using sketch models or small prototypes. However, these tests are often evaluated on metrics of interest to funders who are uninterested in unintended adverse consequences to society.

Some sectors systematically build lower-risk trials into the innovation process [88]. The pharmaceutical industry, which is heavily regulated by government, follows a strict sequence in testing novel drug formulations, first in laboratory conditions on animals, then on a small number of human volunteers, then on a much larger number and variety of human volunteers, and then it performs follow-up studies on results within the broad human population. The electric power industry, which is heavily regulated by government and in some jurisdictions publicly owned, tests innovations on a small scale such as at an individual power plant or substation, before allowing system-wide deployments. For example, “in Denmark’s early phase of developing offshore wind power, utilities were encouraged to experiment with offshore wind, and were allowed to pass on costs to consumers” [89: p. 28].

Other sectors, especially those driven by network effects and that lack tight governmental oversight, such as social media platforms, perform localized trials only to confirm technical performance. New products and services then enter the marketplace at scale. Adverse society-wide impacts sometimes result [90].

The marketplace itself can identify innovations that have unintended adverse consequences, especially if there is a clear way to assign liability for damage caused. The nuclear power industry experienced such disciplining following accidents at Chernobyl and Fukushima, and construction cost overruns at many sites. However, in cases where internalizing the externality proves difficult, governments and social institutions may need to step in to correct failures of market decision making. There is even room for small-scale experimentation in the design of governmental innovation policies [91].

VII. SUGGESTIVE EXAMPLES

This section briefly considers four well-known technological disasters to ask whether the suggested approaches for anticipating adverse unintended consequences would have been feasible at the time of these innovations’ deployments. Is there

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

room for “cautious optimism” [92]? Each case has received prior scholarly attention, so this paper synthesizes previous works rather than interpreting primary sources. As such, reasonable people may differ in their interpretations of the evidence.

TABLE I
ANTICIPATORY TOOLS

Approach	Strengths	Weaknesses	Appropriate Applications
<i>Substantive Methods</i>			
Analogizing	Low barrier to use, draws on history	Limitations unclear beforehand	Early assessment, incremental innovations
Interpolation	Highlights human behavior	Requires fine-grained data	Technologies for everyday living & working
Projection	Formalizes knowledge explicitly & transparently, replicable	High level of effort required	Major long-lived assets, well understood systems
<i>Procedures</i>			
Reflection	Anyone can do it, decentralized	Requires time away from productive activities	Practice where learning can improve future performance
Reasoning	Individual activity often guided by norms	Individuals have different values	Situations involving conflicting public & private interests
Discourse	Broadens information base, diversifies casual stories	Time consuming, potential for miscommunication	Lack of consensus, when contextual knowledge is important
<i>Early Learning Tools</i>			
Observation	Identifies adverse impacts of deployment at scale	Slow and expensive, requires an ongoing resource commitment	Situations in which an adverse effect is hypothesized
Testing	Reduces deployment risk by testing at smaller scale	Delays entry into the marketplace	When empirical support is absent, test sites are available

A. CFC refrigerants

The invention of inert, non-flammable chlorofluorocarbons in 1930 by GM’s Frigidaire division allowed refrigeration to safely scale up to create cold chains for food and medicine, comfortable buildings located in any climate, and many other

beneficial products [93]. Unfortunately, dissipative uses of these same miracle chemicals contributed to global warming and created the Antarctic ozone hole that exposed plants and animals living in the Southern Hemisphere to increased levels of harmful ultraviolet (UV-B) radiation [94]. A global treaty enacted in 1987 phased out the most damaging formulations [95], but a somewhat-diminished ozone hole persists [96].

The following timeline draws on [97]. The industrial R&D to develop CFC refrigerants began in 1928 and resulted in patents by 1930, with industrial-scale production accelerating from the late 1930s and CFCs being measured everywhere in the ambient environment by 1971. The chemical’s ubiquity prompted reflection on the part of the chemical industry which formed a nineteen-company study panel on the ecology of fluorocarbons in 1972. Scientists confirmed existence of a stratospheric ozone layer in 1931 and its role in protecting life on Earth from harmful ultraviolet radiation in 1934, but the link between CFCs and destruction of the stratospheric ozone layer was not identified until 1974. This unintended consequence was confirmed in 1975 using observational data collected from 1957 onward as part of an International Geophysical Year effort. Projections using known chemistry provided the basis in 1974 for Molina and Rowland’s hypothesis, and observations allowed a test of that hypothesis, but only 40 years after widespread deployment of CFCs. Projections, observations, and much discussion facilitated the international agreement in 1987 that phased out these problematic chemicals.

B. Drones

Unmanned aerial vehicles (UAVs) for military purposes date back to French hot-air balloons in 1783, but modern drone technology dates to 1935 when De Havilland deployed a radio-controlled aircraft allowing more realistic target practice training for British military pilots. Development of modern UAVs accelerated with the invention of radio, microprocessors, and geo-spatial positioning systems [98]. UAVs played important roles in Middle Eastern and other conflicts from 1982 onward [99]. In 2006, non-military UAVs were first allowed in U.S. airspace to help respond to the devastation of Hurricane Katrina, and since then a vibrant civilian drone industry has developed. Mass production, standardization, lower costs, and clear regulatory frameworks have encouraged many new applications of drones such as wedding photography, real estate sales, wildlife tracking, and farming [100]. Access to inexpensive versions of this dual-use technology in turn have dramatically empowered those engaged in asymmetrical warfare, as shown in both terrorist attacks and the first year of the underdog Ukrainian response to its invasion by Russia in 2022 [101].

Early thinking about drones relied on analogies to manned aircraft and applications based on remote sensing like what hot air balloons (and later, satellites) could deliver. Even after commercial drones emerged in 2006, many military strategists persisted in viewing military and civilian drones as different and evaluating capabilities according to traditional aircraft-oriented metrics such as payload and speed [102]. However, some engaged in reflection and reasoning to ask whether inexpensive drones would lower the threshold for war [103]. The recent successful use of commercial drones for military purposes by Ukraine eventually prompted Russian military

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

planners to interpolate and adapt [101]. Projections of emerging AI capabilities and potential swarm deployments are leading strategists to anticipate new and highly destructive functionalities for drone fleets fielded by the great powers [104].

C. *Thalidomide*

Chemical Industry Basel (CIBA) in 1952 synthesized the drug Thalidomide, and in 1957 it was acquired by Chemie Grunenthal and marketed as a sedative having fewer side-effects than others available at the time. It was widely used in Japan, Australia, and Europe, but not in the United States because it did not receive FDA approval [105]. Unfortunately, when taken by pregnant women to combat morning sickness, it caused high rates of congenital malformations in their offspring. Thalidomide affected thousands of children. By 1961, it was removed from the market in most countries. Since then, drug companies have increased the rigor with which they test drugs before bringing them to market, including tests for developmental toxicity [106]. Thalidomide has subsequently found use for treating leprosy and a type of bone marrow cancer [107].

The following interprets the rich reporting of [105]. Projection failed because this drug was invented by combining “two rather innocuous compounds...both very safe compounds. When you put them together, we find a very powerful birth-defect causing agent” [105: p. 8]. Analogy also failed, because trials in animal models showed none of the adverse effects that emerged in humans. Procedural tools likely would have helped if they had been implemented. Instead of reflection, the inventors hurried the drug to market and ignored early adverse reports from local physicians. Moral reasoning was tainted by the immediately prior corruption of medical ethics by the Nazi regime. Individual victims and their physicians were not aware that others were also seeing similar symptoms because there were few forums for sharing such information. Instead of conducting experiments with a small number of human subjects to limit risk, the inventors sold their product over the counter to the masses because regulators in Germany did not prevent it [106]. The inventors did not log observations of adverse reports; this was eventually organized by others after the fact.

In the U.S. context, a more consultative procedure was in place in which “safety evaluation involved review of submitted data by a chemist, a pharmacologist, and a medical officer,” none of whom were satisfied with the data received, which delayed approval until the application was withdrawn [108: p. 221]. The medical officer, who had acted while still in her first year on the job, was subsequently “honored with the President’s Award for Distinguished Federal Civilian Service from President John F. Kennedy in 1962” [109: p. 292].

D. *Facebook’s Newsfeed*

Facebook incorporated in 2004 pursuing a mission to “give people the power to share and make the world more open and connected” within ever-expanding online social networks [110]. Advertising provided a path to profitability. Its Newsfeed feature was launched in 2006 to deliver algorithmically selected content to users. It was spectacularly successful in driving growth in the user base and revenues. With growth came a variety of unintended consequences, including a proliferation of pornography, hoaxes, scams, and hate speech in the

Newsfeed. Algorithmic updates addressed these concerns on a piecemeal basis, seeking to balance the potential social harms against the need for revenue. Facebook’s content-serving algorithms and those of other social media platforms are now understood to have spurred political hyper-partisanship, misinformation, and violence at an unprecedented scale in many countries [111]. In response, the company changed its name to Meta and its mission statement to “give people the power to build community and bring the world closer together” [112]. Its founder articulated a new imperative to develop “the social infrastructure for community – for supporting us, for keeping us safe, for informing us, for civic engagement, and for inclusion of all” [113]. It remains to be seen whether this positive outcome emerges.

Algorithmically selected content provided by Facebook and other social media companies has become the major source of news for half of U.S. adults [114]. Since 2005, Americans of all ages have greatly increased their use of social media (from 12% to 84% for ages 18-29, from 2% to 45% for ages 65+) [115, 116]. “Despite a string of controversies and the public’s relatively negative sentiments about aspects of social media, roughly seven-in-ten Americans say they ever use any kind of social media site – a share that has remained relatively stable over the past five years” [117]. The tension in this case is that social media algorithms are achieving an *intended* outcome of holding users’ attention without effectively reducing their unintended consequences because both rely on network effects.

Early analogizing noted that social media, like the printing press, could be expected to induce both positive and negative changes [118]. In 2004, industry actors equated connectedness with social good. Clear evidence that it could also spread social ills came later [119, 120]. Projections rooted in behavioral science were at first inconsistent but have more clearly identified modest negative effects in recent years [121]. Social media companies collect observational data and conduct experiments on an ongoing basis but rarely share the results. Discussion takes place but it is mostly adversarial rather than solutions oriented. Reflection occurs but often those who engage in it must leave their firms. Moral reasoning is becoming enshrined in standards [e.g., 122], although adoption of such standards by the industry is just beginning.

Analogy had value for anticipating consequences of drones and Newsfeed but misled about Thalidomide. Interpolation showed value in the drones case. Projection had value in all cases but was rarely timely. Reflection had value in the CFCs, drones, and Newsfeed cases, and unfortunately seemed not to have been attempted by the manufacturer in the Thalidomide case. Reasoning played a role in the drones and Newsfeed cases. Discussion emerged as a key element in the Thalidomide case. Observation helped in the CFCs and Newsfeeds cases. Testing did not seem to help anticipate unintended consequences in any of cases, although it helped refine products.

The apparent lack of testing for adverse societal consequences in the four brief technological case studies is a noteworthy finding that warrants further investigation. Is it infeasible or just a low priority? The examples suggest that procedural tools may be somewhat more effective than analytical methods in identifying unintended adverse consequences of innovation. This suggestive finding reflects the difficulty of imagining—

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

whether through analogy, projection, or interpolation—surprising outcomes that lie beyond the reach of current scientific understanding. Open discussion seems to help. Very often, it is a game of catching up and remedying adverse outcomes before the innovation becomes widely deployed.

VIII. CONCLUSIONS

Technological innovations sometimes surprise people with unintended adverse consequences when deployed at scale, and current economic and political institutions mostly respond reactively. There are many tools available to aid people in better anticipating these consequences, and they should learn their appropriate applications and use them more widely. However, people are unlikely to avoid adverse consequences all the time, so people will also continue to need to advance their capabilities for timely and effective reaction.

Better anticipation of adverse consequences has utilitarian benefits, but it is also an obligation that innovators owe the rest of society. Individual innovators, organizations that employ them, and nations that host them each can and should perform anticipatory activities. Market forces and political mobilization sometimes help to minimize undesirable outcomes, but they are intrinsically reactive, and interest focused. Anticipation is a difficult task that requires much knowledge and a rich model of cause and effect. But even a simple requirement that invention disclosures include a paragraph of reflection on possible unintended consequences could advance the discussion. Collaborative deliberations can strengthen both the substance of an effort and its impact. Approaches for anticipating adverse consequences are teachable and deserve wider use.

REFERENCES

- [1] W. Shakespeare, *A Midsummer Night's Dream*, Act 3, Scene 2, Line 117 ("Lord, what fools these mortals be!"). (B. Mowat, P. Werstine, eds.). The Folger Shakespeare. Accessed on May 15, 2024 at <https://folger.edu/explore/shakespeares-works/>.
- [2] R.K. Merton, "The Unanticipated Consequences of Purposive Social Action." *American Sociological Review* 1, no. 6 (1936): 894–904. <https://doi.org/10.2307/2084615>.
- [3] E. Roe, "What Are Policy Narratives?: Four Examples and Their Policymaking Implications." In *Narrative Policy Analysis: Theory and Practice*, 34–51. Duke University Press, 1994. <https://doi.org/10.2307/j.ctv11hpgq4.6>.
- [4] C. Geertz, 'Thick Description: Toward an Interpretive Theory of Culture', in his *The Interpretation of Cultures*. New York: Basic Books, 1973: 3–30.
- [5] R.A.W. Rhodes, 'What is decentred analysis?' in his *Narrative Policy Analysis: Cases in Decentred Policy*. London: Palgrave Macmillan, 2019, 1–21. <https://doi.org/10.1007/978-3-319-76635-5>.
- [6] J. Habermas, *Theory of Communicative Action*, trans. Thomas McCarthy, Boston: Beacon Press, 1984.
- [7] C.E. Lindblom, *Inquiry and Change*, New Haven, Yale University Press, 1992.
- [8] R. Vernon, "Unintended Consequences," *Political Theory*, vol. 7, no. 1, 1979, pp. 57–73. Doi: 10.1177/009059177900700104.
- [9] N. Andersson, "Technological Disasters -Towards a Preventive Strategy: A Review," *Tropical Doctor*, vol. 21, no. 1_suppl, pp. 70–81, 1991, doi: 10.1177/004947559102108117.
- [10] E. Scigliano, "10 Technology Disasters," *Technology Review* (June 1, 2002).
- [11] R. Brownsword, E. Scotford and K. Yeung, eds. *The Oxford Handbook of Law, Regulation and Technology*. 2017. Doi: 10.1093/oxfordhb/9780199680832.001.0001.
- [12] B.E. Tonn and D. Stiefel, "Anticipating the Unanticipated-Unintended Consequences of Scientific and Technological Purposive Actions," *World Futures Review*, vol. 11, no. 1, 2019, pp. 19–50. Doi: 10.1177/1946756718789413.
- [13] V. Berthet and V. de Gardelle, "The heuristics-and-biases inventory: An open-source tool to explore individual differences in rationality," *Front Psychol.* 2023 Apr 3;14:1145246. Doi: 10.3389/fpsyg.2023.1145246.
- [14] D. Kahneman, O. Sibony and C.R. Sunstein, *Noise a flaw in human judgement*. William Collins, 2021.
- [15] H.A. Xu, A. Modirshanechi, M.P. Lehmann, W. Gerstner and M.H. Herzog, "Novelty is not surprise: Human exploratory and adaptive behavior in sequential decision-making," *PLoS Comput Biol.* col. 17, no. 6, 2021, pp. e1009070. Doi: 10.1371/journal.pcbi.1009070.
- [16] E.M. Rogers, *Diffusion of innovations* (3rd ed.). New York: Free Press of Glencoe, 1983. ISBN 9780029266502
- [17] T. Hughes-Davies and N. Tamblin, 'Tort Law'. Oxon: Routledge, 2020, pp. 1–19. ISBN 9781138554597.
- [18] R.W. Cobb and C.D. Elder, "The Politics of Agenda-Building: An Alternative Perspective for Modern Democratic Theory," *The Journal of Politics*, vol. 33, no. 4, Nov. 1971. Doi: 10.2307/2128415.
- [19] Y. Benkler, *The Wealth of Networks — How Social Production Transforms Markets and Freedom*. New Haven, CT: Yale University Press, 2006.
- [20] A.D.I. Kramer, J.E. Guillory and J.T. Hancock, "Emotional contagion through social networks," *Proceedings of the National Academy of Sciences*, vol. 111, no. 24, Jun 2014, pp. 8788–8790. Doi: 10.1073/pnas.1320040111.
- [21] T. Aven, Risk assessment and risk management: Review of recent advances on their foundation, *European Journal of Operational Research*, vol. 253, no. 1, 2016, pp. 1–13, Doi: 10.1016/j.ejor.2015.12.023.
- [22] M. Hayenhjelm and J. Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature," *European Journal of Philosophy*, vol. 20, no. S1, 2011, pp. e26–e51. Doi: 10.1111/j.1468-0378.2011.00482.x.
- [23] C.J. Andrews, *Humble Analysis: The Practice of Joint Fact-finding*. Westport CT: Praeger, 2002.
- [24] A. Stirling, "Precaution in the Governance of Technology," In R. Brownsword, E. Scotford and K. Yeung, eds. *The Oxford Handbook of Law, Regulation and Technology*. 2017. Doi: 10.1093/oxfordhb/9780199680832.013.50.
- [25] C. Gollier and N. Treich, "Decision-Making under Scientific Uncertainty: The Economics of the Precautionary Principle," *Journal of Risk and Uncertainty*, vol. 27, no. 1, August 2003, pp. 77–103. Doi: 10.1023/A:1025576823096.
- [26] M. McKelvey and R.J. Saemundsson, "The grey zones of technological innovation: negative unintended consequences as a counterbalance to novelty," *Industry and Innovation*, vol. 28, no. 1, 2021, pp. 79–101. Doi: 10.1080/13662716.2020.1783216.
- [27] J.W. Wetmore, "Amish technology: Reinforcing values and building community," *IEEE Technology & Society Magazine* vol. 26, no. 2, Summer 2007, pp. 10–21.
- [28] H.A. Simon, "From substantive to procedural rationality." In: Kastelein, T.J., Kuipers, S.K., Nijenhuis, W.A., Wagenaar, G.R. (eds) *25 Years of Economic Theory*. Springer, Boston, MA, 1976. https://doi.org/10.1007/978-1-4613-4367-7_6
- [29] P. Dews, Communicative rationality. In *The Routledge Encyclopedia of Philosophy*. Taylor and Francis, 1998. Retrieved 3 Mar. 2024, from <https://www.rep.routledge.com/articles/thematic/communicative-rationality/v-1>. Doi:10.4324/9780415249126-N007-1
- [30] National Research Council, *Understanding Risk: Informing Decisions in a Democratic Society*. Washington, DC: The National Academies Press, 1996. <https://doi.org/10.17226/5138>.
- [31] L.R. Kaplan, M. Farooque, D. Sarewitz, D. Tomblin, "Designing Participatory Technology Assessments: A Reflexive Method for Advancing the Public Role in Science Policy Decision-making," *Technological Forecasting and Social Change*, Volume 171, 2021, 120974, ISSN 0040-1625, <https://doi.org/10.1016/j.techfore.2021.120974>.
- [32] M.G. Morgan and M. Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge, UK: Cambridge University Press, 1990, p. 56.
- [33] W.R. Leonard, "Centennial perspective on human adaptability," *Am J Phys Anthropol.* Vol. 165, 2018, pp. 813–833. Doi: 10.1002/ajpa.23401.
- [34] R.E. Ployhart and P.D. Bliese, "Individual Adaptability (I-ADAPT) Theory: Conceptualizing the Antecedents, Consequences, and Measurement of Individual Differences in Adaptability", Shawn Burke,

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

- C., Pierce, L.G. and Salas, E. (Ed.) *Understanding Adaptability: A Prerequisite for Effective Performance within Complex Environments* (Advances in Human Performance and Cognitive Engineering Research, Vol. 6), Bingley: Emerald Group Publishing Limited, 2006, pp. 3-39. Doi: 10.1016/S1479-3601(05)06001-7.
- [35] National Academies of Sciences, Engineering, and Medicine (NASEM), "Horizon scanning and foresight methods," Ch. 6 (pp. 233-272) in *Safeguarding the Bioeconomy*. Washington, DC: The National Academies Press, 2020. <https://doi.org/10.17226/25525>.
- [36] M. Bazerman and M. Watkins, *Predictable Surprises: The Disasters You Should Have Seen Coming, and How to Prevent Them*. Boston, MA: Harvard Business School Press, 2004, pp. 3-11.
- [37] P. Bordalo, N. Gennaioli and A. Shleifer, "Salience," *Annual Review of Economics*, 2022 vol. 14, no. 1, pp. 521-544, 2022, doi: 10.1146/annurev-economics-051520-011616.
- [38] F. Gaub, INTRODUCTION: BETWEEN PREDICTION AND FORESIGHT. In GLOBAL TRENDS TO 2030: CHALLENGES AND CHOICES FOR EUROPE (pp. 4-6). European Union Institute for Security Studies (EUISS), 2019. <http://www.jstor.org/stable/resrep21147.4>
- [39] P. McKay Illari, F. Russo and J. Williamson, eds., *Causality in the Sciences* (Oxford, 2011; online edn, Oxford Academic, 22 Sept. 2011), doi: 10.1093/acprof:oso/9780199574131.001.0001, accessed 15 Nov. 2023.
- [40] Reddit. 2024. Accessed on March 13, 2024 at https://www.reddit.com/r/sysadmin/comments/bm84zc/what_are_some_good_analogies_you_use_to_explain/
- [41] C.B. Menadue and K.D. Cheer, "Human Culture and Science Fiction: A Review of the Literature, 1980-2016," *SAGE Open* 2017, vol. 7, no. 3, 2017. Doi: 10.1177/21582440177236.
- [42] M. Bar, "The proactive brain: using analogies and associations to generate predictions," *Trends Cogn Sci* (2007)11: 280-289. <https://doi.org/10.1016/j.tics.2007.05.005>.
- [43] S. Baronett, *Logic*. Upper Saddle River, NJ: Pearson Prentice Hall, 2008, pp. 321-325.
- [44] D.J. Staley, "Book Review: Digital Historiography: Analogies," *Journal of the Association for History and Computing*, vol. 4, no. 3, November 2001.
- [45] C. Schwarz-Plaschg, "The Power of Analogies for Imagining and Governing Emerging Technologies," *Nanoethics* 12, 139-153 (2018). <https://doi.org/10.1007/s11569-018-0315-z>
- [46] T. Greenhalgh, D. Swinglehurst, "Studying technology use as social practice: the untapped potential of ethnography," *BMC Med* 9, 45 (2011). <https://doi.org/10.1186/1741-7015-9-45>
- [47] C. Heath, H. Knoblauch, P. Luff, "Technology and social interaction: the emergence of 'workplace studies'," *Br J Sociol.* 2000, 51: 299-320. 10.1080/00071310050030190.
- [48] E. Shove, M. Pantzar and M. Watson, *The Dynamics of Social Practice: Everyday Life and How It Changes*. London: Sage, 2012.
- [49] D. Acemoglu and P. Restrepo, "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy* 2020 128:6, 2188-2244.
- [50] C.J. Andrews, "Restoring legitimacy to the systems approach," *IEEE Technology and Society Magazine*, vol. 19, no. 4, pp. 38-44, Winter 2000/2001.
- [51] A. Sayer, *Method in Social Science: a realist approach*. Hutchinson/Routledge, 1984/1992.
- [52] J.M. Buchanan, "The Public Choice Perspective," *Journal of Public Finance and Public Choice*, vol. 1, no. 1, pp. 7-15, 1983. Doi: 10.1332/251569298X15668907116436.
- [53] A. Giddens, *The constitution of society: Outline of the theory of structuration*. Cambridge: Polity Press, 1984.
- [54] F.W. Geels, "Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study," *Research Policy*, vol. 31, no. 8-9, 2002, pp. 1257-1274. Doi: 10.1016/S0048-7333(02)00062-8.
- [55] J.M. Epstein and R.L. Axtell, *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge, MA: MIT Press, 1996.
- [56] C.J. Andrews, "Rationality in policy decision making," Ch. 12 in F. Fischer, G.J. Miller & M.S. Sidney, eds., *Handbook of Public Policy Analysis*, NY: CRC Press, 2007, pp. 161-71.
- [57] J. Weyant, "Some Contributions of Integrated Assessment Models of Global Climate Change," *Review of Environmental Economics and Policy*, vol. 11, no. 1, 2017, pp. 115-137. Doi: 10.1093/reep/rew018.
- [58] S. Lange, F. Kern, J. Peuckert and T. Santarius, "The Jevons paradox unravelled: A multi-level typology of rebound effects and mechanisms," *Energy Research & Social Science*, vol. 74, 2021, 101982. Doi: 10.1016/j.erss.2021.101982.
- [59] S. Jasanoff, "Technologies of Humility: Citizen Participation in Governing Science," *Minerva* 41, 223-244 (2003). <https://doi.org/10.1023/A:1025557512320>
- [60] D.H. Guston, "Understanding 'anticipatory governance,'" *Soc Stud Sci*, vol. 44, no. 2, 2014, pp. 218-242. Doi: 10.1177/0306312713508669.
- [61] R. Owen, P. Macnaghten and J. Stilgoe, "Responsible research and innovation: From science in society to science for society, with society," *Sci Public Policy*, vol. 39, no. 6, 2012, pp. 751-760. Doi: 10.1093/scipol/scs093.
- [62] S. Urueña, "Responsibility through Anticipation? The 'Future Talk' and the Quest for Plausibility in the Governance of Emerging Technologies," *Nanoethics*, vol. 15, 2021, pp. 271-302. Doi: 10.1007/s11569-021-00408-5.
- [63] J.F. Ahearne, "Three Mile Island and Bhopal: Lessons Learned and Not Learned," pp. 206-220 in National Academy of Engineering. 1986. *Hazards: Technology and Fairness*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/650>.
- [64] D.A. Schön, *The reflective practitioner: how professionals think in action*. New York: Basic Books, 1983. ISBN 978-0465068746.
- [65] M. Cochran-Smith and S.L. Lytle, "Relationships of Knowledge and Practice: Teacher Learning in Communities," *Review of Research in Education*, vol. 24, pp. 249-305, 1999. Doi: 10.2307/1167272.
- [66] G. Bolton, *Reflective Practice: Writing and Professional Development*, 3rd ed. Thousand Oaks, CA: Sage, 2010, p. 7.
- [67] T. Borton, *Reach, touch, and teach: student concerns and process education*. New York: McGraw-Hill, 1970. OCLC 17800
- [68] K. Pretz, "Use a New IEEE Standard to Design a Safer Digital World For Kids," *IEEE Spectrum*, 4 April 2022, <https://spectrum.ieee.org/ieee-standard-protecting-kids-online>,
- [69] J. Parker, "Global Leadership Training and Technology," In V. Wang (Eds.), *Encyclopedia of E-Leadership, Counseling and Training* IGI Global, 2012, pp. 278-287). Doi: 10.4018/978-1-61350-068-2.ch020.
- [70] H.P. Young, "The Evolution of Social Norms," *Annual Review of Economics*, vol. 7, no. 1, 2015, pp. 359-387.
- [71] C.J. Andrews, "From professional ethics to technological citizenship," in *Proceedings of the International Symposium on Electronics and the Environment*, Piscataway, NJ: IEEE, 2006, pp. 45-7.
- [72] IEEE, *IEEE Code of Ethics*. Accessed on November 5, 2021 at <https://www.ieee.org/about/corporate/governance/p7-8.html>.
- [73] C.M. Smith, "Origin and uses of primum non nocere—above all, do no harm!" *J Clin Pharmacol.*, vol. 45, no. 4, Apr 2005, pp. 371-7. Doi: 10.1177/0091270004273680. PMID: 15778417.
- [74] C.J. Andrews, H.C. Jonas, N. Mantell, and R.S. Solomon, Deliberating on statewide energy targets, *Journal of Planning Education and Research*, Summer 2008, 28(1): 6-20. <https://doi.org/10.1177/0739456X08321800>.
- [75] B. Wynne, "Sheepfarming after Chernobyl: A Case Study," *Environment*, vol. 31, no. 2, 1991, pp. 10-15, 33-39.
- [76] J. Forester, *Planning in the Face of Power*. Berkeley, CA: University of California Press, 1989, p. 150.
- [77] A. Deserti, F. Rizzo, & M. Smallman, Experimenting with co-design in STI policy making. *Policy Design and Practice*, (2020)3(2), 135-149. <https://doi.org/10.1080/25741292.2020.1764692>
- [78] J.L. Reynolds, E.B. Kennedy & J. Symons, "If deliberation is the answer, what is the question? Objectives and evaluation of public participation and engagement in science and technology," *Journal of Responsible Innovation*, (2023) 10:1, DOI: 10.1080/23299460.2022.2129543
- [79] A. Buhmann, C. Fieseler, "Towards a deliberative framework for responsible innovation in artificial intelligence," *Technology in Society*, Volume 64, 2021, 101475, <https://doi.org/10.1016/j.techsoc.2020.101475>.
- [80] R. Willis, A. Yuille, P. Bryant, D. McLaren & N. Markusson, "Taking deliberative research online: Lessons from four case studies." *Qualitative Research*, (2023) 23(4), 921-939. <https://doi.org/10.1177/14687941211063483>
- [81] C. Samara, A. Garcia, C. Henry, L. Vallotton, L. Cariolato, J. Desmeules & A. Pinçon, "Safety Surveillance During Drug Development: Comparative Evaluation of Existing Regulations." *Advances in therapy*, (2023) 40(5), 2147-2185. <https://doi.org/10.1007/s12325-023-02492-3>
- [82] W.L. Wallace, *The logic of science in sociology*. Chicago: Aldine-Atherton, 1971.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

- [83] F. El Allaki, M. Bigras-Poulin, P. Michel and A. Ravel, "A population health surveillance theory," *Epidemiol Health*. 2012, 34:e2012007. Doi: 10.4178/epih/e2012007.
- [84] J.M. Diamond, H.A. Latimer II, K.R. Munkittrick, K.W. Thornton, S.M. Bartell and K.A. Kidd, "Prioritizing contaminants of emerging concern for ecological screening assessments," *Environmental Toxicology and Chemistry*, vol. 30, 2011, pp. 2385-2394. Doi: 10.1002/etc.667.
- [85] R. Fuller, P.J. Landrigan, K. Balakrishnan, G. Bathan, S. Bose-O'Reilly, M. Brauer, J. Caravanos, T. Chiles, A. Cohen, L. Corra, M. Cropper, G. Ferraro, J. Hanna, D. Hanrahan, H. Hu, D. Hunter, G. Janata, R. Kupka, B. Lanphear, M. Lichtveld, K. Martin, A. Mustapha, E. Sanchez-Triana, K. Sandilya, L. Schaeffli, J. Shaw, J. Seddon, W. Suk, M.M. Téllez-Rojo, and C. Yan, "Pollution and health: a progress update," *The Lancet Planetary Health*, vol. 6, no. 6, 2022, pp. e535-e547, doi: 10.1016/S2542-5196(22)00090-0.
- [86] R.F. Kizilcec, J. Reich, M. Yeomans, C. Dann, E. Brunskill, G. Lopez, S. Turkey, J.J. Williams, and D. Tingley. Scaling up behavioral science interventions in online education, *Proceedings of the National Academy of Sciences*, Vol. 117, No. 26, 14900-14905. <https://doi.org/10.1073/pnas.1921417117>
- [87] C. Ansell. What is a "democratic experiment"? *Contemporary Pragmatism*. 2012;9(2):159-180. Doi: 10.1163/18758185-90000235.
- [88] Å. Lindholm-Dahlstrand, M. Andersson, & B. Carlsson. Entrepreneurial experimentation: a key function in systems of innovation. *Small Bus Econ* 53, 591-610 (2019). <https://doi.org/10.1007/s11187-018-0072-y>
- [89] M. Grubb, et al. "Induced innovation in energy technologies and systems: a review of evidence and potential implications for CO2 mitigation." *Environmental Research Letters* (2021) 16(4): 043007. DOI 10.1088/1748-9326/abde07
- [90] S.A. Applin, C. Flick, "Facebook's Project Aria indicates problems for responsible innovation when broadly deploying AR and other pervasive technology in the Commons," *Journal of Responsible Technology*, Volume 5, 2021, 100010, <https://doi.org/10.1016/j.jrt.2021.100010>.
- [91] A. Bravo-Biosca, "Experimental Innovation Policy," *Innovation Policy and the Economy* 2020 20:, 191-232. <https://doi.org/10.1086/705644>
- [92] R. Pringle, K. Michael and M. G. Michael, "Unintended Consequences: The Paradox of Technological Potential," in *IEEE Potentials*, vol. 35, no. 5, pp. 7-10, Sept.-Oct. 2016, doi: 10.1109/MPOT.2016.2569672.
- [93] T. Jackson, *Chilled: How Refrigeration Changed the World and Might Do So Again*. Bloomsbury Sigma, 2016. ISBN-13: 978-1472911445.
- [94] J.A. Williams, "This 1920s Inventor Sped Up Climate Change With His Chemical Creations." 2019. Accessed on November 7, 2021 at <https://www.history.com/news/cfcs-leaded-gasoline-inventions-thomas-midgley>.
- [95] United Nations Environmental Programme (UNEP). 2021. "About Montreal Protocol." Accessed on November 7, 2021 at <https://www.unep.org/ozonaction/who-we-are/about-montreal-protocol>.
- [96] National Aeronautics and Space Administration (NASA). "NASA Ozone Watch." 2021. Accessed on November 7, 2021 at <https://ozonewatch.gsfc.nasa.gov/SH.html>.
- [97] S.O. Andersen, M.L. Halberstadt, and N. Borgford-Parnell, "Stratospheric ozone, global warming, and the principle of unintended consequences—An ongoing science and policy success story," *Journal of the Air & Waste Management Association*, vol. 63, no. 6, pp. 607-647, 2013, doi: 10.1080/10962247.2013.791349.
- [98] D. Daly, "A Not-So-Short History of Unmanned Aerial Vehicles (UAV)." 2023. Accessed on November 19, 2023 at <https://consortiq.com/uas-resources/short-history-unmanned-aerial-vehicles-uavs>.
- [99] D.B. Ottaway, "War and missiles," *Washington Post*, June 14, 1982.
- [100] S.J. Höhrvá, J. Soviar and W. Sroka, "Market Analysis of Drones for Civil Use," *LOGI—Scientific Journal on Transport and Logistics*, vol. 14, no. 1, 2023, pp. 55-65. Doi: 10.2478/logi-2023-0006.
- [101] K. Chávez and O. Swed, "Emulating underdogs: Tactical drones in the Russia-Ukraine war," *Contemporary Security Policy*, vol. 44, no. 4, pp. 592-605, 2023. Doi: 10.1080/13523260.2023.2257964.
- [102] M.C. Horowitz, S.E. Kreps and M. Fuhrmann, "Separating Fact from Fiction in the Debate over Drone Proliferation," *International Security*, vol. 41, no. 2, 2016, pp. 7-42. Doi: 10.1162/ISEC_a_00257.
- [103] J.I. Walsh and M. Schulzke, *The Ethics of Drone Strikes: Does Reducing the Cost of Conflict Encourage War?* Strategic Studies Institute, US Army War College, 2015. <https://press.armywarcollege.edu/monographs/442/>.
- [104] B.A. Swett, E.N. Hahn and A.J. Llorens, (2021). "Designing Robots for the Battlefield: State of the Art," In: J.S. von Braun, M. Archer, G.M. Reichbergand M. Sánchez Sorondo, eds., *Robotics, AI, and Humanity*. Cham.: Springer, 2021. Doi: 10.1007/978-3-030-54173-6_11.
- [105] R. Brynner and T. Stephens, *Dark Remedy: The Impact of Thalidomide and Its Revival as a Vital Medicine*. New York: Basic Books, 2001.
- [106] C. Bennett, *History of Thalidomide*. 2020. Accessed on November 7, 2021 at <https://www.news-medical.net/health/History-of-Thalidomide.aspx>.
- [107] V. Pannikar, "The Return of Thalidomide: New Uses and Renewed Concerns," *WHO Pharmaceuticals Newsletter*, no. 2, 2003, pp. 11-12.
- [108] F.O. Kelsey, "Thalidomide update: Regulatory aspects." *Teratology* (1988) 38: 221-226. <https://doi.org/10.1002/tera.1420380305>
- [109] W. Rehman, L.M. Arfons, H.M. Lazarus, "The rise, fall and subsequent triumph of thalidomide: lessons learned in drug development." *Ther Adv Hematol*. 2011 Oct;2(5):291-308. Doi: 10.1177/2040620711413165. PMID: 23556097; PMCID: PMC3573415.
- [110] M.J. Coren, "Facebook's global expansion no longer has its mission statement standing in the way," *Quartz* (22 June 2017). Accessed on November 7, 2021 at <https://qz.com/1012461/facebook-changes-its-mission-statement-from-ing-its-mission-statement-from-sharing-making-the-world-more-open-and-connected-to-build-community-and-bring-the-world-closer-together/>.
- [111] J. Lepore, "Facebook's Broken Vows," *The New Yorker Magazine*, 2 August 2021.
- [112] Meta. Mission Statement. 2021. Accessed on November 7, 2021 at <https://about.facebook.com/company-info/>.
- [113] M. Zuckerberg, Building Global Community (5 May 2021 update). Accessed on November 7, 2021 at <https://www.facebook.com/notes/3707971095882612/>.
- [114] J. Liedke and L. Wang, "Social Media and News Factsheet." Pew Research, 2023. Accessed on December 3, 2023 at <https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/>.
- [115] M. Faverio, 2022. "Share of those 65 and older who are tech users has increased in the last decade." Pew Research Center, 2022. Accessed on December 3, 2023 at <https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/>.
- [116] A. Perrin, "Social Networking Usage: 2005-2015." Pew Research Center. October 2015. Accessed on December 3, 2023 at <http://www.pewinternet.org/2015/10/08/2015/Social-Networking-Usage-2005-2015/>
- [117] Pew Research Center, April 2021, "Social Media Use in 2021"
- [118] J.A. Dewar, "The information age and the printing press: Looking backward to see ahead." RAND, 1998. Accessed on December 3, 2023 at <https://www.rand.org/content/dam/rand/pubs/papers/2005/P8014.pdf>.
- [119] J. Daniels, "Race and Racism in Internet Studies: A Review and Critique." *New Media & Society*, vol. 15, no. 5, pp. 695-719, 2013.
- [120] A. Matamoros-Fernández and J. Farkas, "Racism, Hate Speech, and Social Media: A Systematic Review and Critique," *Television & New Media*, vol. 22, no. 2, 2021, pp. 205-224. Doi: 10.1177/1527476420982230.
- [121] E. Kross, P. Verduyn, G. Sheppes, C.K. Costello, J. Jonides and O. Ybarra, "Social Media and Well-Being: Pitfalls, Progress, and Next Steps," *Trends in Cognitive Sciences*, vol. 25, no. 1, 2021, pp. 55-66, doi: 10.1016/j.tics.2020.10.005.
- [122] IEEE Standard for an Age Appropriate Digital Services Framework Based on the 5Rights Principles for Children, IEEE 2089-2021.



Clinton J. Andrews (Life Senior Member, IEEE) received the BS degree in engineering with Honors from Brown University, Providence, RI, USA in 1978; the MS degree in technology and policy (Sigma Xi) from the Massachusetts Institute of Technology, Cambridge, MA, USA in 1985; and the PhD degree in urban and regional planning from MIT in 1990.

He initially worked in the private sector and at Princeton University, and since 1997 he has been at Rutgers University, New Brunswick, NJ, USA, serving as a professor, center director, and the associate dean for research at the EJ Bloustein

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

School of Planning and Public Policy. His work addresses how technological and environmental changes affect urban life. His books include *Humble Analysis: The Practice of Joint Fact-finding* (Praeger 2002), *Regulating Regional Power Systems* (IEEE Press/Westport 1995), and *Industrial Ecology and Global Change* (Cambridge University Press, 1994).

Dr. Andrews has served on several editorial boards and as chief co-editor of the *Journal of Planning Education and Research*. He remains a member of the Trend Scouting community for the American Planning Association. He is a member of the American Institute of Certified Planners, a licensed Professional Engineer, and a Fellow of AAAS. He is a past president of the IEEE Society on Social Implications of Technology.