

The Corporate Villainy of Tom Sawyer:

Manufactured Scarcity, the Ethics of Attention, and the Literary Prehistory of Dark Patterns

Travis Gilly

ORCID: 0009-0007-2954-6313

Real Safety AI Foundation

Illinois, United States

Working paper, June 2026 (v1)

ABSTRACT

The whitewashing scene in Mark Twain's *The Adventures of Tom Sawyer* (1876) is conventionally received in one of two ways: as a comic illustration of the malleability of value, the reading made canonical in behavioral economics by Ariely, Loewenstein, and Prelec (2006), or as a celebration of the boyish individualist, the reading carried into popular culture by works such as Rush's 1981 song of the same name. This paper proposes a third reading. It argues that the scene is the earliest structurally complete literary rendering of the design logic now catalogued under the heading of dark patterns: manufactured scarcity, engineered desire, and the extraction of value from a manipulated audience. On this reading Tom is not a charming rascal but the prototype of the engagement designer who optimizes a system against the people inside it. The paper distinguishes two ethics of gamification, the earned, in which an activity is genuinely made worth doing and the designer participates in it, and the engineered, in which desire for access to a hollow good is manufactured and the designer abstains, and it proposes the designer's own abstention, the refusal to consume the product, as a diagnostic of the engineered kind. Locating the reading within the ethics of attention and persuasive design (Fogg, 2003; Verbeek, 2011; Williams, 2018) and the dark patterns literature (Brignull, 2010; Mathur et al., 2019), it contends that the fence is not a quaint antecedent but a working model of contemporary attention capture, and that the uniformly admiring reception of Tom is itself evidence of how thoroughly the culture has naturalized the extractor as hero. It closes on the irony that Twain, who coined the term Gilded Age and named the capitalist as the oppressor, narrated the con as cleverness and never marked it as harm.

Keywords: philosophy of technology, ethics of attention, persuasive design, dark patterns, gamification, manufactured scarcity, behavioral economics, Mark Twain

I. INTRODUCTION

A boy is sentenced to whitewash a fence. He does not want to do it, and at first he tries to buy his way out, offering another boy his small treasures to take the brush. That fails. Then something turns over in him. He stops behaving as though the chore is a punishment and begins behaving as though it is a rare privilege, an art that few can perform and that he is reluctant to share. Other boys, free on a summer Saturday, watch, then want, then pay. By the end of the morning the fence has three coats of whitewash, Tom Sawyer is, in Twain's phrase, literally rolling in wealth, and he has not lifted the brush since the first mark.

The scene is among the most quoted in American literature, and it has been read in two dominant registers. In the first, the register of behavioral economics, it is a parable about the construction of value: a demonstration that people often do not possess a stable, pre-existing sense of whether an experience is pleasant or unpleasant, so that the same act can be coded as drudgery or as delight depending on the cues that frame it (Ariely, Loewenstein, & Prelec, 2006). In the second, the register of popular culture, it is a celebration of the figure of Tom himself, the resourceful and unconforming American boy, an archetype durable enough that nearly a century later a rock band could borrow the name for an anthem about the individualist whose mind, in the song's refrain, is not for rent (Lee, Lifeson, Peart, & Dubois, 1981).

Both readings are defensible. Neither is complete, because neither asks the question that the scene answers most precisely. The question is not whether value can be constructed, which the behavioral economists have established, nor whether Tom is likeable, which the culture has assumed. The question is what, exactly, Tom does to the other boys, and whether the thing he does has a name in the present.

It does. This paper argues that the whitewashing scene is the earliest clean literary model of the design logic now studied under the heading of dark patterns and, more broadly, of persuasive and attention capturing design (Brignull, 2010; Fogg, 2003; Mathur et al., 2019). Tom does not make the work better. He manufactures scarcity around access to it, engineers covetousness, and extracts value from boys who walk away with nothing but a turn at a chore. Read this way, Tom is not the hero of the scene. He is its villain, and the scene is a small and complete schematic of the extraction economy that a century and a half of technology would eventually industrialize.

The paper makes three claims. First, descriptively, that the mechanism in the fence scene is manufactured scarcity and engineered desire rather than genuine engagement, and that the standard behavioral reading, by treating the scene as a neutral curiosity about valuation, has left this normative structure unexamined. Second, conceptually, that gamification divides into two ethics, an earned ethic in which an activity is made authentically worth doing and the designer partakes in the value created, and an engineered ethic in which desire for a hollow good is manufactured and the designer abstains, and that the designer's own abstention furnishes a portable diagnostic of the second kind. Third, genealogically, that the fence is a structurally complete instance of the modern dark pattern, so that the literature on attention capture has a literary prehistory it has not claimed, and that the uniformly flattering reception of Tom is itself evidence of how successfully the culture naturalizes the extractor as a figure of admiration.

A review of the literature locates no prior treatment of the whitewashing scene as a model of dark pattern design or within the ethics of attention. The scene's existing scholarly uses run in two other directions. In behavioral economics, and in the law and economics that followed it, the fence is a case of the construction of value and of motivation crowding, the reframing of an activity in a way that alters how it is valued (Ariely, Loewenstein, & Prelec, 2006; Lepper, Greene, & Nisbett, 1973; Kim, 2021). In literary criticism, the novels have been read through game theory, the colonial question, and, above all, the moral suspicion of Tom set against Huck (Marx, 1953). Each component the present reading draws on is therefore well established, the dark patterns taxonomy included (Brignull, 2010; Mathur et al., 2019). The contribution is the joining, not the parts. A claim assembled from established materials is sturdier, not weaker, for having its supports in view.

II. TWO RECEPTIONS OF THE FENCE

A. The Construction of Value

The most influential modern treatment of the fence is not literary but economic. In "Tom Sawyer and the construction of value," Ariely, Loewenstein, and Prelec (2006) take the scene as the epigraph for a challenge to the standard economic assumption that agents know their own tastes. Across a series of studies they show that people can be induced to treat the same experience as something to be paid for or something to be paid to avoid, depending on an initial and essentially

arbitrary cue. The scene earns its place at the head of the paper because it dramatizes exactly this: the whitewashing is fixed, but its valence is not, and Tom flips the valence by reframing the activity from obligation to privilege. The same scene anchors a related strand in behavioral law and economics, where it serves as an illustration of motivation crowding, the way an external frame or incentive can displace or manufacture the willingness to act (Lepper, Greene, & Nisbett, 1973; Kim, 2021).

This reading is correct as far as it goes, and it goes only as far as description. The construction of value, in the behavioral account, is a fact about cognition. It is neither good nor bad; it is simply a way the mind can be moved. The framing leaves the normative question untouched, because the discipline that took up the scene is in the business of explaining behavior, not evaluating designs. What the behavioral reading does not say, and was never trying to say, is that the construction of value is also a tool, that tools are used by someone on someone, and that the someone in the fence scene is a child being relieved of his property by another child who has discovered how to do it. The mechanism is named. The transaction is left in the dark.

B. The Admirable Rascal

If the economists neutralized the scene, the culture flattered it. Across its reception Tom has been the lovable scamp, the embodiment of a particular American freedom: improvisational, authority skeptical, a little larcenous in a way the reader is invited to enjoy. The whitewashing is told as a triumph, and the triumph is Tom's. We are positioned with the con artist, never with his marks.

The clearest modern monument to this reception is Rush's 1981 song "Tom Sawyer" (Lee, Lifeson, Peart, & Dubois, 1981). The song is not a retelling of any episode; it borrows the name as shorthand for a modern individualist, the free spirited figure whose defining boast is that his mind is not for rent to any god or government. The band's lyricist described the source poem as a portrait of a modern day rebel striding through the world wide eyed and purposeful. The point for present purposes is not the song's quality, which is considerable, but its direction. Even a deliberate reinvention of the character reached for Tom as an emblem of autonomy and refusal. The reach is revealing, and it contains an irony the song does not register: the anthem of the mind that is not for rent takes its name from the one figure in the source text whose entire achievement is renting out the minds and labor of others, and pocketing the rent. The free man of the chorus is, in the founding scene, the landlord of the con.

Two receptions, then, and a shared blind spot. The economists saw a mechanism and declined to judge it. The culture saw a hero and declined to look. Between them they have kept the scene in view for a hundred and fifty years without anyone turning the fence around to read what is written on the back.

III. WHAT TOM ACTUALLY DOES

A close reading of the episode reveals a structure more deliberate than either reception admits, and three features of that structure carry the argument.

The first is that Tom's intention is established before his inspiration, and it is pure escape. Set to whitewash the fence as punishment, his first move is not to reframe the work but to offload it by purchase: he offers another boy his marble and other small valuables to take the brush. The offer is interrupted, but it has already done its evidentiary work. It tells us what Tom wants, which is to be free of the chore, and it tells us what would count, for him, as enough, which is simply not having to do it. This matters because it gives the scene a measurable baseline against which everything that follows can be judged.

The second feature is the mechanism Tom actually uses once the inspiration arrives, which is not genuine engagement but manufactured scarcity. He does not make whitewashing intrinsically rewarding, and he does not believe

it has become so. He makes access to it appear scarce and exclusive, an art that only he can perform acceptably and that he is unwilling to surrender. Twain states the operative law in the narrator's own voice: Tom had discovered, without knowing it, that in order to make a person covet a thing it is only necessary to make the thing difficult to attain (Twain, 1876, ch. 2). That is not a principle of motivation. It is a principle of inducement. The boys do not come to love painting; they come to want a turn at a thing that has been made artificially hard to get, which is a different psychological event with a different ethics. Twain even supplies the moral coloring himself, describing the marks not as customers but as innocents, and Tom, resting between transactions, as planning the slaughter of more of them.

The third feature is the satiation point, and it is the hinge of the villainy. Tom's stated need, escape from the chore, is fully met at the first transaction, the moment Ben Rogers offers an apple for the privilege of painting. At that instant someone else is doing the work, for free to Tom, and Tom is released. The thing he had been willing to pay for, a minute earlier, he has now obtained at no cost. The story could end there with his purpose satisfied and his conduct innocent. It does not end there. Tom does not stop at free. He reverses the polarity of the original transaction, from paying to escape to being paid to escape, and keeps collecting: the apple, then a kite, then a dead rat on a string, then a small museum of boyhood treasures, until the fence has three coats and Tom is rolling in wealth. The extraction is precisely measurable because Tom's own earlier offer to buy his way out marks where enough was, and his conduct sails past it.

These three features compose a single moral shape. Legitimate self interest has a satiation point: the need is met and the taking stops. What the scene depicts is the treatment of that point as a floor rather than a ceiling, the conversion of a met need into a standing opportunity to extract. And it adds a final detail that will prove diagnostic. Across the entire episode, after the inspiration lands, Tom never lifts the brush. He is the one figure present who does not paint the fence. The designer of the desire does not partake of the thing he has made desirable, because he knows there is nothing there to partake of.

IV. TWO ETHICS OF GAMIFICATION

The fence scene is often cited, loosely, as an early case of gamification, the turning of an unpleasant task into something game like and therefore willingly done. The citation is right that a transformation occurs and wrong about what kind it is, and the error is worth correcting because it conceals a distinction on which a great deal of design ethics depends. Gamification is not one practice with one moral valence. It divides into two, and the fence belongs decisively to the worse of them.

A. The Earned Ethic

In the first kind, which may be called earned gamification, the activity itself is genuinely reconstructed so that performing it becomes intrinsically rewarding. The transformation is real, the value created is real, and crucially the designer believes in it and shares in it. Consider a familiar honest case: a group of children set to clear a field, who invent among themselves a game out of the labor, race and rule and mock ceremony, such that the work is completed and they finish wanting to keep playing. Nothing has been manufactured at anyone's expense. No scarcity has been engineered, no mark has been left poorer, and the people who designed the game are inside it, playing. The same structure describes good teaching, in which a subject is made worth learning rather than merely made to seem urgent, and the serious tradition in educational and prosocial game design, which aims to make genuinely valuable activity genuinely engaging (the broad ambition surveyed under headings such as games for change and games for learning). In the earned ethic, the designer is a participant in the value and would consume the product without coercion.

B. The Engineered Ethic

In the second kind, which may be called engineered gamification, no real value is added to the activity. What is added is desire for access to it, produced by manufactured scarcity, exclusivity, urgency, competition, and the other tactics that the behavioral and design literatures have catalogued. The activity remains hollow; only the wanting is new. The player is not a participant but a mark, and the designer does not partake, because the designer knows there is nothing to partake of. This is Tom's kind. He adds no value to whitewashing. He adds covetousness to its access, collects the proceeds, and never paints.

The distinction yields a diagnostic that is the central conceptual contribution of this paper. The tell of engineered gamification is the designer's own abstention. The honest gamifier uses the thing built: the teacher who loves the subject, the field game's inventors who play it, the designer who would willingly be a user. The engineered gamifier refuses the product, because the product is for marks. Tom never lifts the brush. The modern instance of the same tell is well documented and widely remarked: the practice, among some who design habit forming consumer technologies, of restricting those products in their own households and for their own children. Whatever the merits of any particular report, the pattern names the diagnostic precisely. When the makers of a thing will not let the thing near the people they love, the thing is built for slaughter, in Twain's word, and not for use. Abstention is the confession.

This diagnostic does real work, because it answers the most common objection to any critique of gamification, the objection that everyone gamifies and that the critique therefore proves too much. It does not prove too much, because it does not condemn gamification. It condemns one of the two ethics into which gamification divides, and it supplies a test, the test of the designer's own participation, by which the two can be told apart in any concrete case.

V. THE LITERARY PREHISTORY OF DARK PATTERNS

The tactics by which Tom manufactures desire are not idiosyncratic to a boy and a fence. They are, item for item, the tactics that contemporary human computer interaction research has named, measured, and condemned as dark patterns, the deceptive interface designs that steer people toward choices that serve the designer at the user's expense (Brignull, 2010; Mathur et al., 2019; Mathur, Kshirsagar, & Mayer, 2021).

Set the fence beside the taxonomy. Tom's central device, making access to the brush appear limited and difficult, is the pattern the literature calls artificial or manufactured scarcity. His insistence that the privilege may be withdrawn at any moment, and that the present chance is fleeting, is the pattern of false urgency, the countdown and the limited time offer that exploit scarcity and loss aversion to rush a decision. His staging of a queue of eager boys, each one's interest validating the next one's, is social proof. And the whole performance is built to redirect attention and reframe value, which is the function that recent work isolates under the heading of attention capture deceptive design (Monge Roffarello, Lukoff, & De Russis, 2023). The fence is not a loose analogue of one modern pattern. It is a compact instance of several at once, assembled into a working system, in 1876.

The genealogical claim follows. The behavioral economists identified the cognitive mechanism the scene exploits, the constructability of value. The interaction researchers identified the design tactics the scene deploys, the patterns of scarcity, urgency, and social proof. What neither has done is connect the literary template to the lineage of extraction it anticipates, and that connection is the point. The whitewashed fence is the earliest structurally complete portrait we have of the practice that the attention economy would later automate and scale: the manufacture of desire for access to a good whose value has been engineered rather than created, and the harvesting of the resulting traffic.

The modern terminus of the lineage is concrete and, by now, adjudicated. In 2022 the Federal Trade Commission resolved actions against the maker of *Fortnite* for a combined \$520 million, of which \$245 million addressed dark patterns, deceptive interface designs that charged users, including children, for purchases they did not intend to make,

and \$275 million addressed violations of children's privacy law (Federal Trade Commission, 2022). The currency of that economy is, in part, the limited time cosmetic item, the skin placed behind a manufactured countdown so that the desire to acquire it is produced by its impending unavailability rather than by any feature of the thing itself. The dead rat on a string and the limited time skin are the same move. The triviality of the object is not a defense; it is the camouflage. A con that traffics in apples and kittens reads as comedy, and reading as comedy is exactly how it has remained, for a hundred and fifty years, a beloved scene rather than a recognized template of harm.

VI. THE ETHICS OF ATTENTION

To say that the fence anticipates dark pattern design is to locate it, finally, within a particular tradition of the philosophy of technology, the tradition concerned with persuasive design and the ethics of attention. Three commitments of that tradition make the placement precise.

The first is that artifacts are not neutral. Winner's argument that artifacts have politics established that a built thing can embody and enforce arrangements of power and value, that design is a way of legislating conduct (Winner, 1980). The fence, as Tom rebuilds it, is such an artifact. It is no longer only a surface to be painted; it is a small machine for converting other children's desire into Tom's wealth, and it carries that politics whether or not anyone names it.

The second is that technologies mediate, in Verbeek's sense, the desires and actions of those who encounter them: they do not merely serve preexisting wants but help to shape what is wanted and how (Verbeek, 2011), a lineage that runs back to the founding of persuasive technology as a field of study (Fogg, 2003). Tom does not satisfy a desire the boys arrive with. He produces the desire, by mediation, through the staging of scarcity. The wanting is an artifact of the design.

The third, and the normative center of the matter, is supplied by the ethics of attention. Williams argues that the competition for human attention has become the defining design problem of the age, and that the deep harm of attention capturing design is not merely that it wastes time but that it erodes the capacity for self determination, the ability to want what we want to want and to act on it (Williams, 2018). This is the register in which the engineered ethic of gamification finally shows its cost. The harm of Tom's fence is not that the boys spend a Saturday painting. It is that their wanting has been authored by someone whose interest is opposed to theirs, and that, inside the manufactured scarcity, painting the fence has been made the rational thing to want. Engineered gamification does not only capture attention. It optimizes against the capacity to disengage, by arranging the situation so that the desire it manufactures is the reasonable response to the conditions it has set.

This last point carries a consequence that the legal and policy literatures are only beginning to register, and that can be stated here as a premise the present argument supports rather than a claim it defends in full. A design that optimizes against the capacity to disengage falls hardest on those for whom disengagement is already the most costly, the populations whose attentional self regulation is, by disposition or condition, least able to resist an environment engineered to defeat it. To engineer against the capacity to disengage is, for those populations, to engineer against the disability itself. The fence is a tidy emblem of this, because the boy with the least power to resist a manufactured scarcity is the one most likely to surrender his treasures for a turn at the brush.

VII. THE AUTHOR AND THE CON

There is a final turn, and it belongs to Twain. The author who wrote the cleanest miniature of the extraction economy was, in his public life, that economy's most quotable critic. Twain coined the term by which the era of that economy is still known, in the title of his 1873 novel *The Gilded Age*, a satire of speculation and political corruption written

with Charles Dudley Warner (Twain & Warner, 1873). He was a committed friend of organized labor, who praised the Knights of Labor and, in an 1886 address, asked who the oppressors were and answered that they were the few, the king and the capitalist and their overseers, set against the many who make the bread the idle eat (Twain, 1886/1958). In his last decade he served as a vice president of the Anti Imperialist League. The man held, and said, that concentrated economic power preys on those who work.

And yet the same man wrote the whitewashing scene and narrated it as a triumph. He told the story of a small extraction, complete in every part, in the admiring key of a clever boy's discovery of a great law of human action, and he did not mark the law as predatory. There is an honesty constraint to observe here, and it is important to observe it precisely, because the temptation is to overclaim. We cannot read Twain's mind, and we should not assert that he failed to see what his scene contains. The claim available from the text is narrower and stronger than a claim about the author's awareness. It is that the scene is narrated as cleverness to be enjoyed rather than as extraction to be condemned, and that this admiring narration sits in unremarked tension with everything Twain argued elsewhere about the few who prey on the many. Whether he missed the symptom or simply declined to flag it, the contradiction is on the page, and the reception history records that no one collected it.

The irony deepens at the level of biography, and here too the figure is doubled rather than simple. Twain was not only a critic of the speculative economy; he was an enthusiastic participant in it, a speculator in mining stock, a founder of a publishing house, an investor who poured money into inventions, and a man whose own ventures eventually cost him a great deal (Twain, 1990). The author who named the age of extraction and called the capitalist the oppressor was himself, in his ledgers, a little bit Tom Sawyer, forever chasing the apple. The fork the scene contains, between the honest maker who participates and the extractor who manufactures and abstains, ran through its author's own life.

One clarification protects the reading from a natural misunderstanding. To call the scene ahead of its time is to misplace the achievement. The mechanism was not ahead of its time at all; it was the defining engine of Twain's present, the manufactured shortage and the watered stock and the engineered scarcity of the Gilded Age he had himself diagnosed. What is ahead of its time is the diagnosis, not the disease. Twain rendered his own moment with such economy that the single image still functions, a century and a half later, as an accurate schematic of an extraction economy he could not have foreseen in its digital form. He did not predict the future. He saw the present so clearly that the picture still reads as a diagram of ours.

VIII. OBJECTIONS

Three objections deserve direct answer.

The first holds that the reading crushes a children's joke under an apparatus it cannot bear, that a comic episode in a boys' book is being made to carry the weight of an economic and ethical theory it was never built for. The answer is in the text. Twain himself does not leave the scene at the level of the particular; his narrator universalizes it into a great law of human action and into a general account of the difference between work and play. The scaling to adult and institutional cases is licensed by Twain's own framing, not imposed against it. And the apparent triviality, the apple and the dead rat, is not a reason to dismiss the reading but the very feature that explains the scene's long immunity. A con conducted in trinkets reads as charm, and reading as charm is how a complete model of extraction has passed, for generations, as a delight.

The second objection holds that there is nothing new in finding Tom suspect, since a substantial critical tradition already reads him as the morally compromised figure of the novels, most influentially in the long argument over the cruelty of the closing sequence of *Huckleberry Finn*, where Tom's elaborate game prolongs Jim's captivity for the sake of Tom's own entertainment (Marx, 1953). This is correct, and it is an ally rather than a rival. The tradition establishes

that Tom can be read against the grain of his own charm. What it has not done is give the suspicion an economic and design theoretic content, reading the fence specifically as the template of manufactured scarcity and engineered desire. The novelty claimed here is not that Tom is flawed, which is old, but the frame, which is the genealogy of extraction.

The third objection holds that the argument proves too much, that since reframing and incentive design are everywhere in teaching, parenting, and ordinary persuasion, a critique that catches the fence catches all of them. This is answered by the distinction of Section IV. The argument does not condemn gamification or reframing as such. It condemns the engineered ethic specifically, the manufacture of desire for a hollow good at the expense of a mark, and it supplies the diagnostic, the designer's own abstention, by which the engineered ethic is told apart from the earned one in any particular case. The field game the children invent for themselves and the fence Tom stages for his marks are the same mechanism turned to opposite ethics, and the test of who participates in the value distinguishes them cleanly.

IX. CONCLUSION

The whitewashed fence is the first dark pattern. It is a complete and compact model of the practice that the attention economy would later automate: the manufacture of desire for access to a good whose value has been engineered rather than created, and the extraction of value from the people in whom that desire is produced. Read so, the scene yields three things the standard receptions have missed. It yields a distinction, between the earned ethic of gamification, in which value is genuinely created and the designer partakes, and the engineered ethic, in which desire is manufactured and the designer abstains. It yields a diagnostic, the designer's own abstention, the refusal to consume the product, as the tell of the engineered kind. And it yields a piece of evidence about ourselves, in the form of a hundred and fifty years of admiring reception that kept polishing the figure of Tom and never once turned the fence around, a reception so thorough that the culture's modern emblem of the unrentable mind is the very boy whose achievement was to rent the minds of others.

The practical recommendation is small and follows directly. Design ethics should treat manufactured scarcity and engineered desire as objects of default suspicion rather than as neutral tools of engagement, and the honest tradition of gamification, the games for learning and games for change that aim to make valuable activity genuinely engaging, should be willing to submit to the abstention test, to show that its designers would let their own children near the thing they have built. Tom would not let anyone near the brush except to pay for it. That, and not his cleverness, is the thing the scene has been trying to tell us, in the key of comedy, since 1876.

ACKNOWLEDGMENTS

The author discloses the use of speech-to-text software (Wispr Flow) and Anthropic’s Claude as assistive technology for a diagnosed neurodevelopmental condition. All intellectual contributions, research questions, theoretical arguments, methodological decisions, and conclusions are solely the author’s own.

Correspondence concerning this article should be addressed to Travis Gilly, Real Safety AI Foundation, Illinois, United States. Email: t.gilly@ai-literacy-labs.org.

REFERENCES

- Ariely, D., Loewenstein, G., & Prelec, D. (2006). Tom Sawyer and the construction of value. *Journal of Economic Behavior & Organization*, 60(1), 1–10.
- Brignull, H. (2010). *Deceptive design* (formerly Dark Patterns). Retrieved from <https://www.deceptive.design>
- Federal Trade Commission. (2022, December 19). *Fortnite video game maker Epic Games to pay more than \$500 million over FTC allegations of privacy violations and unwanted charges* [Press release]. Retrieved from <https://www.ftc.gov>
- Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann.
- Kim, H. (2021). Can mandating corporate social responsibility backfire? *Journal of Empirical Legal Studies*, 18(1), 189–251.
- Lee, G., Lifeson, A., Peart, N., & Dubois, P. (1981). Tom Sawyer [Song]. On *Moving Pictures*. Anthem/Mercury.
- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children’s intrinsic interest with extrinsic reward: A test of the “overjustification” hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129–137.
- Marx, L. (1953). Mr. Eliot, Mr. Trilling, and *Huckleberry Finn*. *The American Scholar*, 22(4), 423–440.
- Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1–32.
- Mathur, A., Kshirsagar, M., & Mayer, J. (2021). What makes a dark pattern... dark? Design attributes, normative considerations, and measurement methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–18).
- Monge Roffarello, A., Lukoff, K., & De Russis, L. (2023). Defining and identifying attention capture deceptive designs in digital interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–19).
- Twain, M. (1876). *The Adventures of Tom Sawyer*. American Publishing Company.
- Twain, M. (1958). The new dynasty (1886). In P. S. Foner, *Mark Twain: Social critic*. International Publishers.
- Twain, M. (1990). *The autobiography of Mark Twain*. (C. Neider, Ed.). Harper & Row.
- Twain, M., & Warner, C. D. (1873). *The Gilded Age: A tale of today*. American Publishing Company.
- Verbeek, P.-P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.

Williams, J. (2018). *Stand out of our light: Freedom and resistance in the attention economy*. Cambridge University Press.

Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136.