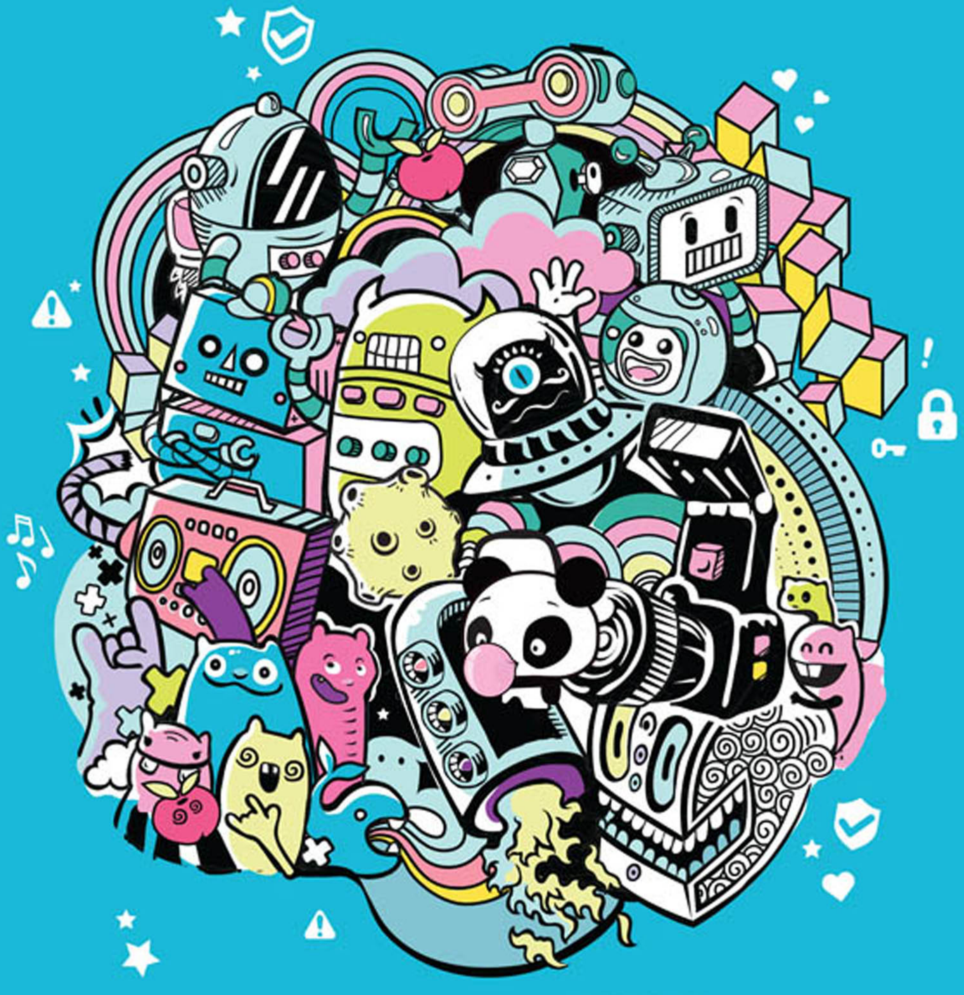
Algorithmic Rights and Protections for Children



EDITED BY

Mizuko Ito, Remy Cross, Karthik Dinakar, and Candice Odgers

Algorithmic Rights and Protections for Children

Algorithmic Rights and Protections for Children

edited by Mizuko Ito, Remy Cross, Karthik Dinakar, and Candice Odgers

The MIT Press

Cambridge, Massachusetts | London, England

The open access edition of this book was made possible by generous funding from the MIT Libraries.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Ito, Mizuko editor. | Cross, Remy, editor. | Dinakar, Karthik, editor. | Odgers, Candice L., editor.
Title: Algorithmic rights and protections for children / edited by Mizuko Ito, Remy Cross, Karthik Dinakar, Candice Odgers.
Description: Cambridge, Massachusetts : The MIT Press, 2023. | Includes bibliographical references and index.
Identifiers: LCCN 2022034631 (print) | LCCN 2022034632 (ebook) | ISBN 9780262545488 (paperback) | ISBN 9780262374323 (epub) | ISBN 9780262374316 (pdf)
Subjects: LCSH: Computers and children—United States. | Internet and children—United States. | Computer algorithms—Social aspects—United States. | Data privacy—United States. | Computer literacy—United States. | Internet—Safety measures. | Children's rights—United States.
Classification: LCC QA76.9.C659 A44 2023 (print) | LCC QA76.9.C659 (ebook) | DDC 004.083—dc23/eng/20221108
LC record available at https://lccn.loc.gov/2022034631
LC ebook record available at https://lccn.loc.gov/2022034632

Contents

Acknowledgments

We are grateful for all the contributors to this collection over the course of both its online publication as a Work in Progress on MIT Press Open, through several rounds of revisions, before the print and online publication as a completed volume. The development and revisions of the essays for this volume happened during the tumultuous ups and downs of the COVID pandemic from 2020 to early 2022. We have so much appreciation for all the contributors to this volume for their tenacity and goodwill, as well as their intellectual brilliance through a lengthy pathway to publication during tough times.

Susan Buckley at the MIT Press deserves special thanks for guiding this volume through both the MIT Press Open and its final publication. We would also like to thank Samantha Bates and Joi Ito, who helped lead the initial call for proposals that launched this volume. Jenna Abrams at the Connected Learning Lab at the University of California, Irvine offered a keen editorial eye as well as careful and caring stewardship of every essay in the volume through many stages of submission, revision, and refinement.

Part I Perspectives

1
Introduction

Algorithmic Rights and Protections for Children

Mizuko Ito, Remy Cross, Karthik Dinakar, and Candice Odgers

One in three internet users worldwide is a child (Livingstone, 2015), and what children see and experience online is increasingly shaped by algorithms. Yet the dominant platforms of the online world have not been constructed with the needs and interests of children in mind. Children represent an especially marginalized and vulnerable population exposed to high levels of poverty and inequality, while being dependent on adults to advocate for their interests and structure their experiences. In 2023, as we are still recovering from a pandemic that has made us even more reliant on digital platforms, society is struggling to rein in the power of big tech and elevate the needs of marginalized groups. This tension is particularly acute when it comes to balancing opportunities and risks for children in online spaces.

Social media, educational technologies, and networked games have been a lifeline to social connection and learning during the COVID pandemic. As schools began to reopen after the first wave of the pandemic, a third of children in the US, particularly students of color, said they would prefer to continue to learn online and not return to the classroom (Schwartz et al., 2020). Some parents who once discouraged their children from playing Fortnite and Roblox (Flake, 2021) now see these platforms as an essential social outlet (Kelly, 2021). In tandem with this growing reliance on digital platforms, concerns over children's digital privacy, safety, rights, and inequality are also mounting (Barassi, 2020; Livingstone et al., 2018; Zuboff, 2020). Whether it is search results (Noble, 2018), video recommendations on YouTube, or assessing student learning (Williamson, 2017), algorithms are beginning to gain influence on young people's well-being, learning, and future opportunity. Because young people are a uniquely vulnerable group, supporting healthy online engagement for children is the tip of the spear for regulation of digital platforms, and one of the thorniest arenas for balancing protection and rights.

Despite the important role that children's protections and rights play in debates of the social impacts and responsibility of tech platforms, issues unique to children have not been a significant focus of debates over artificial intelligence (AI) and ethics. Some notable exceptions include UNICEF's AI for Children project (UNICEF, n.d.), the work of organizations such as the Family Online Safety Institute (FOSI, n.d.), Common Sense Media (Common Sense Media, n.d.), the 5Rights Foundation (5Rights, n.d.), and the UN Committee on the Rights of the Child's General Comment 25 (United Nations, n.d.), outlining children's rights in digital spaces. A small but growing body of work on digital parenting and children's experiences with algorithms seeks to inform this debate (see, e.g., Barassi, 2020; Lenhart & Owens, 2021; Livingstone & Blum-Ross, 2020; Livingstone et al., 2018). This collection of essays builds on this momentum, providing perspectives, frameworks, and research for understanding diverse children's evolving relationships with algorithms, and how caregivers, educators, policy-makers, and other adult stakeholders might shape these relationships in productive ways. We introduce the collection by outlining three cross-cutting concerns: (1) the relationship between algorithms, culture, and society; (2) the unique needs and positionality of children; and (3) inequality in children's risks and opportunities.

Algorithms, Culture, and Society

Despite the often novel nature of algorithms, big data, and AI, our existing frameworks for understanding the relationship between technology, culture, and society are as relevant as ever. Science and technology studies scholars have insisted that we look at how technologies are *shaped by* our existing cultural biases and institutionalized practices, and also how they *shape* culture and society (see, e.g., Hine, 2016; MacKenzie & Wajcman, 1985; Bijker et al., 2012). The time is ripe for critical scrutiny of how algorithms are shaped by and reflect historical inequities, problematic assumptions, and institutionalized power. We also need solution-oriented scholarship and design thinking that considers how these technologies can be shaped to be more equitable and serve the needs and interests of children. This volume includes work that critically analyzes how algorithms reflect existing structures and biases, as well as work centered on designing and reshaping technology to serve children.

Algorithms and their impacts are inseparable from the institutional dynamics that children encounter in schools, community-based organizations, and families and with commercial entertainment and communication industries. We need a critical understanding of how technology grows out of the specific social, institutional, and cultural contexts that define and constrain diverse forms of contemporary childhoods. For example, today's "revolutionary" educational technologies can reflect entrenched interests as well as outdated assumptions about learning and automating instruction (Losh, 2014; Reich, 2020; Watters, 2021). In this collection, Paulo and Izidoro Blikstein describe how today's technology solutionist rhetoric around automating instruction and assessment has deep roots in early generations of educational technology. Maureen Mauk (this volume) considers a growing burden on parents to manage and monitor media—what she describes as "responsibilization"—that has roots in the nineteenth century.

Even as new technologies grow out of and are shaped by entrenched structures and assumptions, how they are being developed, institutionalized, and taken up in everyday life are very much under negotiation and public debate. The nature of these negotiations differs depending on which stage the technology is at in innovation, spread, and societal adoption and adaptation.

Some technologies, such as relational robots for children (see Boulicault et al., this volume), are just emerging from the research lab. Others, such as algorithms for monitoring and predicting youth violence (see Patton et al., this volume), are just beginning to be rolled out and are encountering resistance from stakeholders. Still other technologies and platforms, such as online video and search, voice assistants, and learning management systems, are already "domesticated" (Haddon, 2011) and in widespread use (see O'Byrne et al.; Druga et al.; Manago et al., all in this volume).

Many contributors to this collection have focused on how we might productively shape and reshape emerging technologies to empower children and be more responsive to their needs. Marion Boulicault, Milo Phillips-Brown, Jacqueline M. Kory-Westlund, Stephanie Nguyen, and Cynthia Breazeal are building and testing relational robots in partnership with young children and educators. Their contribution challenges established assumptions about authenticity and child-robot relationships, suggesting ways of designing relationships that support and honor the unique perspectives of young children who experience robots differently from adults. Drawing from her experiences as a school-based technology integrationist, Michelle Ciccone suggests ways that educators can critically evaluate classroom digital tools as one step toward supporting student digital literacies. Sayamindu Dasgupta and Benjamin Mako Hill surface the ways in which young people themselves are understanding, interrogating, and critiquing algorithmic systems in the context of the Scratch online community. They derive a set of design principles for algorithmic literacy and engagement from these observations. These and other essays in this volume elevate the voices and agency of varied stakeholders in reshaping and defining algorithmic technologies with which children engage.

These complexities and nuances demand a multifaceted, interdisciplinary, and international dialogue. The diverse perspectives represented in this collection, though far from comprehensive, offer a sampling of the range of viewpoints and frameworks that need to be at the table during this moment of rupture and debate, when practices and policies are in flux on varied fronts. Contributors represent fields as wide-ranging as social work, robotics, educational research, instructional design, design research, and media studies. While the agency and influence of scholars and innovators may be limited in an arena dominated by big tech and high-stakes global political wrangling, we hope that interdisciplinary coalitions of researchers and innovators can continue to raise issues and offer framings that are grounded in longstanding field and disciplinary wisdom, as suggested by our contributors.

Children's Perspectives and Needs

AI challenges our assumptions, most obviously about what counts as intelligence, and the boundaries between humans and machines. Perhaps less obviously, AI also challenges us to reconsider assumptions about childhood culture, what is "age appropriate," and the balance between rights and protections for children. Negotiations over media and technology have long been a site of intergenerational struggle. Whether novels, television, video games, or today's social media, the "new" media of the day have offered an arena for young people to exercise agency and develop new cultural forms, often provoking concern from parents and moral panics writ large (Livingstone & Blum-Ross, 2020; Ito et al., 2019; Jenkins, 1998; Seiter, 1995). The rapid incursion of digital, interactive, mobile, and networked media in young people's lives since the nineties has been a particularly complex and fraught arena for navigating the tension between rights and protections for children. Media and tech companies, and the algorithms that pervade online spaces, are now powerful players in the everyday negotiations over even young children's engagement with knowledge, media, and social networks.

How we protect and empower children in relation to digital technology is made more complex by their changing needs as they grow older. As digital technology moves into the early years, children have the tools to make independent media choices earlier than prior generations. In a 2020 survey, one-third of US parents with children under 12 say their child interacts with a voice-activated assistant, and the same proportion of parents say their child began engaging with a smartphone before the age of 5 (Auxier et al., 2020). In another 2020 survey, 95 percent of parents of children aged 5–8 said that their children watch online videos, and that the children themselves are most likely to select what they watch, rather than the parent (Rideout & Robb, 2020). Developmental science suggests that early adolescence (aged 10–14) is a particularly important time for caregivers to support growing independence and range in media choices, and scaffold first steps into social online spaces. Older adolescents' engagements with technology more closely resemble those of adults', peer-to-peer dynamics are more salient, and teens chafe at overly restrictive parental monitoring and control (Odgers & Robb, 2020, pp. 35–37).

This growing agency and early access to online communication and content has challenged caregivers' and educators' ability to keep up, monitor, and regulate. As parents fret over screen time, stranger danger, and privacy concerns, children's perspectives and interests must also be at the table. Childhood studies scholars have noted how adults tend to view children as "becomings" rather than full "beings," arguing for deferred gratification and preparation for an adult future. Adults often fail to recognize children's unique social and moral perspectives, rights, and interests in the present (James & Prout, 2014; Qvortrup, 2009; Qvortrup et al., 2009). This divergence of interests manifests in everyday family struggles over screen time, as well as in policy frameworks that focus on rights versus protection of children. Researchers have noted how these power dynamics and conflicts over screen time can be more harmful to adolescents' mental health than screen time itself (Mauk, this volume; Odgers & Robb, 2020). In educational settings, the datification and "personalization" of learning and outcomes has become a high-stakes battlefield over issues of learner agency, privacy, and control (Watters, 2021; Williamson, 2017).

Many of the essays in this volume are centered on children's voices and viewpoints, suggesting ways of shaping our algorithmic futures based on these perspectives. Nicholas Santer, Adriana Manago, Allison Starks, and Stephanie Reich conducted a survey of 11–14-year-olds on their views of digital privacy, finding that they are more concerned about privacy from peers and family members than corporate surveillance. Stefania Druga, Jason Yip, Michael Preston, and Devin Dillon involved both children and parents in codesigning an AI literacy framework, informed by their findings that children perceive AI bias differently from adults. Four media literacy scholars—Ian O'Byrne, Kristen Turner, Kathleen A. Paciga, and Elizabeth Stevens—describe conversations with their children about digital technologies and strategies they developed together to productively shape their engagement with online algorithms. These and other contributions help center our consideration of algorithmic rights and protections on young people and their changing perspectives as they grow older (see also this volume: Boulicault and Phillips-Brown et al.; Dasgupta & Hill; Vasudevan).

Unequal Childhoods

The unequal power dynamics between children and adults are critical factors in considering algorithmic rights and protections for children; inequality between different populations of children is equally important. Safiya Noble (2018) opens her book *Algorithms of Oppression* with her experience of googling "Black girls" in hopes of finding interesting content for her stepdaughter and nieces, only to discover pornography featuring Black girls as the first search result. Algorithmic biases and inequalities that pervade the adult world are doubly damaging for marginalized children. We now have a growing literature on the harm that AI and algorithms can cause when they reproduce the assumptions and structural inequalities of the dominant culture (e.g., Benjamin, 2019; Brayne, 2020) but still relatively little work that looks at the impacts on unequal childhoods.

Too often, research and public discourse makes generalizations about the experiences of "kids these days" that ignores the experiences of oppressed and marginalized youth. Essays in this volume build on a budding body of research that examines how social media, digital games, and learning technologies reflect and reinforce unequal childhoods. This includes work on how inequality in children's experiences with technology differ across national contexts (e.g., Global Kids Online, n.d.), as well as within them. For example, scholars have examined how LGBTQ (Cho, 2017; 2015), neurodiverse (Alper, 2017; Ringland, 2019), and BIPOC (Watkins, 2010; Tanksley, 2019) youth experience and engage with social media in unique ways. Also relevant is research on how educational technologies intersect with long-standing inequities in our education systems (Rafalow, 2020; Williamson, 2017; Livingstone & Sefton-Green, 2016; Watkins et al., 2018).

These themes of difference and inequality recur throughout the essays in this collection. Desmond Patton, Siva Mathiyazhagan, and Aviv Y. Landau consider differences in children's experiences with technologies and the state in India, Israel, and the United States. Veena Vasudevan takes a close ethnographic look at the experiences of youth of color and personalized educational technologies. Sayamindu Dasgupta and Benjamin Mako Hill describe how young coders debate the potentially discriminatory implications of the code they are writing and deploying online. Too often, public debates over children, teens, and technology fail to fully recognize the diversity of youth experiences, risks, and benefits, leading to one-size-fits-all policies that take White and middle-class childhoods in the Global North as the baseline. The essays in this volume seek to nuance this picture through deeper dives into the experiences of diverse children in specific contexts.

This Collection

Understanding children's algorithmic rights and protections requires multidisciplinary and cross-sector viewpoints and synthesis, given the range of institutional settings where children encounter algorithms, and the unique forms of inequality and risks that children encounter while growing up. This collection of essays, which began as responses to a call for papers in the *Journal of Design and Science*, represents a variety of viewpoints, fields, and disciplinary voices in two genres. "Perspectives" are shorter conceptual pieces that share a unique viewpoint or apply a framework from a particular field of discipline to the topic at hand. "Full Papers" are longer contributions that report on empirical or design research. The essays offer critical and provocative analysis, frameworks for understanding, and practical approaches for how to productively engage with emerging technologies as designers, educators, and parents. We hope that this range of voices and contributions will foster more dialogue, creative thinking, and coalition building at this unique but critical nexus of children, algorithms, care, and social justice.

References

Alper, M. (2017). *Giving voice*. Amsterdam University Press.

Auxier, B., Anderson, M., Perrin, A., & Turner, E. (2020, July 28). *Parenting children in the age of screens*. Pew Research Center. https://www.pewresearch.org/internet/2020/07/28/childrens-engagement-with-digital-devices-screen-time/

Barassi, V. (2020). *Child data citizen: How tech companies are profiling us from before birth*. The MIT Press.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim Code* (1st ed.). Polity.

Bijker, W. E., Hughes, T. P., Pinch, T., & Douglas, D. G. (2012). *The social construction of technological systems: New directions in the sociology and history of technology* (Anniversary ed.). The MIT Press.

Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing* (1st ed.). Oxford University Press.

Cho, A. (2015). *Sensuous participation: Queer youth of color, affect, and social media* [Doctoral dissertation, University of Texas at Austin]. Texas ScholarWorks.

Cho, A. (2017). Default publicness: Queer youth of color, social media, and being outed by the machine. *New Media & Society*, *20*(9), 3183–3200. https://doi.org/10.1177/1461444817744784

Common Sense Media. (n.d.). Advocating for kids' well-being in the digital age. Retrieved May 21, 2021, from https://www.commonsensemedia.org/kids-action

Family Online Safety Institute. (n.d.). *Family Online Safety Institute*. Retrieved May 21, 2021, from https://www.fosi.org/

5Rights Foundation. *About us*. (n.d.). Retrieved May 21, 2021, from https://5rightsfoundation.com/about-us/

Flake, E. (2021, March 10). My kid sold her soul to Roblox. *The New York Times*. https://www.nytimes.com/2020/09/23/parenting/roblox-screentime.html?te=1&nl=nyt-parenting&emc=edit_ptg_20210120

Global Kids Online. (n.d.). *Global kids online*. Retrieved May 21, 2021, from http://globalkidsonline.net/

Haddon, L. (2011). Domestication analysis, objects of study, and the centrality of technologies in everyday life. *Canadian Journal of Communication*, *36*(2), 311–323. ISSN 1499–6642

Hine, C. (2016). From virtual ethnography to the embedded, embodied, everyday internet. In L. Hjorth, H. Horst, A. Galloway, & G. Bell (Eds.), *The Routledge companion to digital ethnography*. Routledge. https://doi.org/10.4324/9781315673974

Ito, M., Baumer, S., Bittanti, M., boyd, d., & Cody, R. (2019). *Hanging out, messing around, and geeking out: Kids living and learning with new media* (Tenth Anniversary ed.). The MIT Press.

James, A., & Prout, A. (2014). *Constructing and reconstructing childhood: Contemporary issues in the sociological study of childhood (Routledge Education Classic Edition)* (3rd ed.). Routledge.

Jenkins, H. (1998). *The children's culture reader*. NYU Press.

Kelly, H. (2021, March 20). They laughed, they cried, they killed monsters: How friendships thrived in video games during the pandemic. *The Washington Post*. https://www.washingtonpost.com/technology/2021/03/20/video-games-pandemic-friends/?mc_cid=124f9628c9&mc_eid=517c563d02

Lenhart, A., & Owens, K. (2021, May). *The unseen teen: The challenges of building healthy tech for young people*. Data & Society. https://datasociety.net/wp-content/uploads/2021/05/The-Unseen-Teen-.pdf

Livingstone, S. (2015, November 2). *One in three: Internet governance and children's rights*. LSE Media Policy Project Blog. https://blogs.lse.ac.uk/medialse/2015/11/02/one-in-three-internet-governance-and-childrens-rights/

Livingstone, S., & Blum-Ross, A. (2020). *Parenting for a digital future: How hopes and fears about technology shape children's lives* (1st ed.). Oxford University Press.

Livingstone, S., & Sefton-Green, J. (2016). *The class: living and learning in the digital age (Connected youth and digital futures, 1)*. NYU Press.

Livingstone, S., Stoilova, M., & Nandagiri, R. (2018). *Children's data and privacy online: Growing up in a digital age: An evidence review*. LSE Media and Communications. https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review-final.pdf

Losh, E. (2014). *The war on learning*. The MIT Press. https://doi.org/10.7551/mitpress/9861.001.0001

MacKenzie, D., & Wajcman, J. (Eds.). (1985). *The social shaping of technology: How the refrigerator got its hum*. Milton Keynes, Open University Press.

Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Odgers, C., & Robb, M. B. (2020). *Tweens, teens, tech, and mental health: Coming of age in an increasingly digital, uncertain, and unequal world, 2020*. Common Sense Media.

Qvortrup, J. (2009). Are children human beings or human becomings? A critical assessment of outcome thinking. *Rivista Internazionale di Scienze Sociali, 117*(3/4), 631–653. http://www.jstor.org/stable/41625246

Qvortrup, J., Corsaro, W., & Honig, M. (2009). *The Palgrave handbook of childhood studies*. Palgrave Macmillan. https://doi.org/10.1111/j.1099-0860.2011.00420.x

Rafalow, M. H. (2020). *Digital divisions: How schools create inequality in the tech era* (1st ed.). University of Chicago Press.

Reich, J. (2020). *Failure to disrupt: Why technology alone can't transform education*. Harvard University Press.

Rideout, V., & Robb, M. B. (2020). *The Common Sense census: Media use by kids age zero to eight, 2020*. Common Sense Media. https://www.commonsensemedia.org/sites/default/files/research/report/2020_zero_to_eight_census_final_web.pdf

Ringland, K. E. (2019). A place to play: The (dis)abled embodied experience for autistic children in online spaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (paper 288, pp. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300518

Schwartz, H. L., Grant, D., Diliberti, M. K., Hunter, G. P., & Setodji, C. M. (2020). *Remote learning is here to stay: Results from the first American School District Panel Survey*. Creative Commons Attribution 4.0 International Public License, 2020. Rand Corporation. https://www.rand.org/pubs/research_reports/RRA956-1.html

Seiter, E. (1995). *Sold separately: Children and parents in consumer culture (Communications, media, and culture series)*. Rutgers University Press.

Tanksley, T. (2019). *Race, education and #BlackLivesMatter: How social media activism shapes the educational experiences of Black college-age women* [Doctoral dissertation, University of California, Los Angeles]. eScholarship.

UNICEF. (n.d.). *AI for children*. UNICEF Office of Global Insight & Policy. Retrieved May 21, 2021, from https://www.unicef.org/globalinsight/featured-projects/ai-children

United Nations Human Rights Office of the High Commissioner. (1990). *Convention on the Rights of the Child*. United Nations. https://www.ohchr.org/en/professionalinterest/pages/crc.aspx

Watkins, C. S. (2010). *The young and the digital: What the migration to social network sites, games, and anytime, anywhere media means for our future*. Beacon Press.

Watkins, C. S., Cho, A., Lombana-Bermudez, A., Shaw, V., Vickery, J. R., & Weinzimmer, L. (2018). *The digital edge: How Black and Latino youth navigate digital inequality (Connected youth and digital futures, 4)*. NYU Press.

Watters, A. (2021). *Teaching machines: The history of personalized learning*. The MIT Press.

Williamson, B. (2017). *Big data in education: The digital future of learning, policy and practice* (1st ed.). Sage Publications.

Zuboff, S. (2020). *The age of surveillance capitalism: The fight for a human future at the new frontier of power* (Illustrated ed.). PublicAffairs.

2

Algorithmic Literacies

K–12 Realities and Possibilities

Michelle Ciccone

Uneven Digital Literacies Set the Stage

When we speak of "algorithmic rights and protections of children," of course what happens within schools is of particular consequence. Conversations about children and technology have evolved from a hyperfocus on risks to a more balanced consideration of opportunity enabled by digital technologies (Gasser & Cortesi, 2017), with access—to devices, skills, and literacies—being central to this evolving conversation. K–12 schools continue to thread this needle, establishing policies and practices that allow students to access the opportunities afforded by digital technologies while ideally minimizing the risk of negative consequences. This means that, via educational experiences in the classroom, young people may exercise their right of access to and use of powerful algorithmic-driven technologies, but in so doing may become overly exposed to the sorting and tracking mechanisms enabled by these very same technologies. What's more, schools and districts may have protectionist policies in place that aim to reduce student data exposure to unknown and problematic algorithmic systems, but it is not uncommon for these policies to be inconsistently enforced or widely misunderstood by school staff. Critical data literacies (Pangrazio & Selwyn, 2019) help educators and students grapple with the implications of algorithmic systems and their place within those systems so that they might be able to take more informed action, but developing these critical data literacies can be hampered by a lack of general background knowledge, confidence, or interest on the part of school staff, including classroom teachers.

As a school-based K–12 technology integrationist, I see a wide range in algorithmic literacies in educators and students. I have come to understand that the reasons for this are numerous, but let us start with a fundamental issue: the varying levels of basic digital skills of many educators. To be clear, this is not unique to the K–12 educator workforce: adults in the United States in general do not fare well on measurements of "problem solving in technology-rich environments" as compared to international counterparts (Vanek, 2017), adding up to adults in many sectors in the US demonstrating limited digital literacy skills. And though remote schooling during the COVID-19 pandemic has certainly accelerated the development of deeper and more nuanced digital skills and literacies for many teachers, thinking *with* technology remains difficult for many adults in the workforce, so much so that any device or software becomes a "black box" to the user. It is no wonder, then, that algorithms and the literacies that would help make sense of them feel impenetrable to those who might still be gaining fluency in basic digital skills. This, then, begs the question: Where in K–12 schooling do students develop algorithmic literacies?

Finding a Place for Algorithms in the Curriculum

Within the K–12 curriculum, mention of algorithms is often housed solely within computer science courses, in part because there is a sense that algorithmic literacy is a highly technical and specialized skill. For example, in the 2016 Massachusetts Digital Literacy and Computer Science (DLCS) curriculum frameworks and standards, algorithmic literacy is firmly situated within the Computational Thinking strand, which itself is repeatedly tied to programming first and foremost. Algorithmic literacy is of course central to the education of a future computer scientist, but given the ubiquity of algorithms and algorithmic systems, is it not important for the user of these technologies to exercise some degree of algorithmic literacy as well? What's more, tying algorithmic literacy primarily to computer science does not touch all students, as we know that gender and racial inequalities persist in the enrollment patterns for computer science courses (Code.org Advocacy Coalition).

Algorithmic systems are important to study not just to learn their architectures and internal machinations so that we can build them ourselves as computer scientists; they are also worth studying because these algorithmic systems have wide-reaching consequences for society. Within the Massachusetts DLCS standards, there is indeed a Computing and Society strand of standards that asks for curriculum that considers "the beneficial and harmful effects of computing innovations" (p. 47). But algorithms are not mentioned here, even though these harmful effects cannot be truly understood divorced from discussion of the "coded inequities" (Benjamin, 2019) (re)inscribed by algorithmic systems. Examples of these algorithmically coded inequities abound, and engaging case studies for the classroom that consider any number of these real-life examples can be integrated into any content area in the K–12 curriculum. We miss these opportunities if we see algorithmic systems as only the purview of computer scientists.

For educators who may have limited algorithmic background knowledge, where to start may seem daunting, and it is currently difficult to find ready-made classroom materials that provide examples of what algorithmic literacy across the curriculum can look like. In my work, I have developed just such curriculum, in the hopes of inspiring curiosity about algorithms in both students and teachers. As one example, I have developed and implemented curriculum for eighth- and ninth-grade research contexts that asks students to grapple with the algorithmic bias evidenced by Google Images search results for “unprofessional hair” and “three black kids” versus “three white kids” (Noble, 2018). We discuss: For what do you turn to Google Images, and what process do you take to evaluate these search results? How do Google Images search results impact what you believe to be true or “standard”? What do initiatives like World White Web (information about this initiative archived at https://johannaburai.com/World-White-Web) demonstrate that everyday users like us can do to impact the Google Images search algorithm? What is our responsibility as users of these algorithmic systems versus the responsibility of developers to fine-tune these algorithms?

Time and again I rediscover that young people are interested in learning about biased algorithmic systems, and I have found that my teacher colleagues across content areas are genuinely interested as well. The motivation to empower-with-knowledge baked into algorithmic literacy education aligns well with the motivations that lead educators to the teaching profession in the first place: to make a difference to society and in a child’s life (Menzies et al., 2015). Yes, the teaching profession is full of idealists.

But the truth is that engaging in algorithmic literacy is inherently political work, and that can be scary for some educators, especially in this cultural moment that sees anti-racist and antibias curriculum questioned and even banned by school boards and state legislatures across the United States. To truly examine algorithmic systems, students must consider who gets to define and categorize, why certain entities and not others get to do this influential work, and how power structures get reinscribed via these algorithmic systems. In many cases, the developers of these algorithms are not explicitly setting out to create racist, misogynist, and bigoted systems, and biased results are more a reflection of larger societal prejudices. Algorithmic literacy curriculum demands that students grapple with intent of inventors versus impact of inventions, which can be a challenging conversation for some educators to have with students. The desire (and need) of some educators to appear apolitical in the classroom is then coupled with the reality that many educators may not be convinced that it is within their content-area purview to incorporate algorithmic topics in their curriculum due to narrow definitions of what belongs where within a traditional K–12 curriculum. This confluence of ideologies too often leads to missed opportunities. Preparing students to live fully informed lives in an algorithmic culture (Striphas, 2015) becomes someone else's job, and often, in practice, winds up being taken up by no one.

Developing the Algorithmic Literacies of Educators

So how might we demonstrate that algorithmic literacies can (and should) be developed wherever we encounter algorithms? (Which, in fact, is everywhere.) Research from Choi, Cristol, and Gimbert (2018) suggests, "Before promoting advanced levels of digital citizenship, teachers need to successfully achieve online activities in democratic and varied ways" (p. 154). For our purposes here, this means that if we hope that educators engage students in critically examining algorithmic systems, then we must engage educators in this work first.

One place to start could be in developing educators' ability to evaluate classroom digital tools. Generally speaking, K–12 teachers have some degree of autonomy in choosing supplementary curriculum materials, which increasingly include algorithmically driven apps and websites (e.g., skill-building software, digital product creation tools, digital portfolios) in which student work may be connected to personally identifiable information. Teachers may learn about new digital tools at educational conferences and start using these tools when they return to their classrooms, perhaps seeing the workshop presenter as a sufficient vetter of the digital tool. In many districts, there is indeed a process by which digital tools must be approved for classroom use: there is often a person in charge of reading privacy policies and terms of service documents who will determine whether the tool in question complies with the district's policies. There are even organizations such as the Student Data Privacy Consortium (https://privacy.a4l.org/) that can facilitate negotiated contracts with edtech vendors. Permission slips are often sent home to parents and guardians that link to the digital tool's terms of service and request consent for their child to use the given tool. But, as mentioned, even when decision makers have thoughtful, values-driven policies that guide their decisions about which digital tools comply with district policies (and unfortunately this is not always the case), those values are not always shared or understood by the wider district community. The result is uninformed consent at all points in the decision-making process, and a sense from classroom teachers that they could never engage in this analysis on their own.

What if, instead, through ongoing professional development, we equipped and empowered teachers themselves to make these evaluative decisions with their students' algorithmic best interests in mind? What if we start this training in teacher preparation programs, so that teachers enter the classroom practiced in asking critical, technoethical questions (Krutka et al., 2019) when they encounter a supplementary classroom technology? We can equip teachers to ask of edtech products:[1] How does this tool determine what is relevant, correct, or worth knowing? Does this line up with my own educational philosophy? How are my classroom practices being reshaped to suit the algorithmically driven processes of this tool? What data are being collected by this tool, and what is being done with the data? What types of predictions are being made, and do those predictions line up with my pedagogical goals? These questions can even be explored with the help of students so that a technoethical audit (Krutka et al., 2021) becomes a shared learning activity.

To be sure, some districts already equip and empower teachers to ask these questions and make these determinations. And, unfortunately, in at least as many districts, classroom teachers do not have the autonomy to authentically evaluate digital tools introduced by administrator-level decision makers at all. But no matter our starting point, if we have any hope of developing an algorithmically literate generation, one able to exercise and demand their own algorithmic rights, it is clear that we cannot ignore the algorithmic literacies of the educators who teach them today.

Note

1. Questions inspired by Kris Shaffer's *Data, Code, Ethics* seminar for the Digital Pedagogy Lab and Gillespie (2013).

References

Benjamin, R. (2019). *Race after technology*. Polity Press.

Choi, M., Cristol, D., & Gimbert, B. (2018). Teachers as digital citizens: The influence of individual backgrounds, internet use and psychological characteristics on teachers' levels of digital citizenship. *Computers & Education*, *121*, 143–161.

Code.org Advocacy Coalition. 2018. *State of computer science education*. Retrieved August 24, 2022, from https://advocacy.code.org/stateofcs

Gasser, U. & Cortesi, S. (2017). Children's rights and digital technologies: Introduction to the discourse and some meta-observations. In M. Ruck, M. Peterson-Badali, & M. Freeman (Eds.), *Handbook of children's rights: Global and multidisciplinary perspectives* (pp. 417–436). Routledge.

Gillespie, T. (2013). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society*.

University Press Scholarship. https://doi.org/10.7551/mitpress/9780262525374.003.0009

Krutka, D. G., Heath, M. K., & Willet, K. B. S. (2019). Foregrounding technoethics: Toward critical perspectives in technology and teacher education. *Journal of Technology and Teacher Education*, *27*(4), 555–574.

Krutka, D. G., Smits, R. M., & Willhelm, T. A. (2021). Don't be evil: Should we use Google in schools? *TechTrends*, *65*(4), 421–431.

Massachusetts Department of Elementary and Secondary Education. (2016). *Massachusetts Digital Literacy and Computer Science (DLCS) curriculum framework*. Retrieved August 24, 2022, from https://www.doe.mass.edu/frameworks/dlcs.pdf

Menzies, L., Parameshwaran, M., Trethewey, A., Shaw, B., Baars, S., & Chiong, C. (2015). *Why Teach?* LKMco. https://doi.org/10.13140/RG.2.2.12227.86567

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Pangrazio, L., & Selwyn, N. (2019). "'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data." *New Media & Society*, *21*(2), 419–437.

Shaffer, K. (2019, August 5–9). *Data, code, ethics—DPL 2019* [Seminar at Digital Pedagogy Lab]. Seminar description and materials retrieved from https://dpl.pushpullfork.com/

Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, *18*(4–5), 395–412. https://doi.org/10.1177/1367549415577392

Vanek, J. B. (2017). *Using the PIAAC framework for problem solving in technology-rich environments to guide instruction: An introduction for adult educators*. Program for the International Assessment of Adult Competencies. Retrieved August 24, 2022, from https://static1.squarespace.com/static/51bb74b8e4b0139570ddf020/t/589a3d3c1e5b6cd7b42cddcb/1486503229769/PSTRE_Guide_Vanek_2017.pdf

3

Co-Constructing Digital Futures

Parents and Children Becoming Thoughtful, Connected, and Critical Users of Digital Technologies

W. Ian O'Byrne, Kristen Hawley Turner, Kathleen A. Paciga, and Elizabeth Y. Stevens

"What's an algorithm?" This question is not one that most children would ask their parents. But when the parent is a literacy/technology researcher, interesting conversations about computers and their functions seep into the home.

Algorithms are increasingly part of everyday life, and children, as they engage on digital devices, are affected by programs written by companies whose primary goal is to sell content and products. These same companies promote apps that capture the attention of youth of all ages, often under the guise of entertainment, education, or connecting individuals in a community.

As literacy/technology researchers, we understand that children live in and shape a connected world where they have the ability to consume and create literally at their fingertips. We care deeply about preparing them to be lifelong learners with the skills they need to access, analyze, evaluate, create, and participate through digital technologies (Ito et al., 2013).

We are also parents who must navigate the realities of a digital world: every time our children log into an app on a device they are using at school, they leave a data trail. We know they engage in the affordances of digital technologies often through the price of their privacy (Berson & Berson, 2006). At the same time, we know that developing digital literacy includes understanding that algorithms drive users to particular content (Burrell, 2016).

Historically, parents have adopted a range of strategies when thinking about children, screens, and technology. In their study of screens in the homes of children in the UK, Livingstone and Blum-Ross (2020) identified three genres of parenting practices, framed by particular values and beliefs, around the use of new digital media, tools, and technologies in their homes and in their children's lives. The first of these Livingstone and Blum-Ross call *embrace*. Here, parents welcome new technologies and harness them for some sort of specific use. The second parenting practice, which Livingstone and Blum-Ross call *balance*, is marked by parents "encouraging some digital practices and not others, often ad hoc, weighing opportunities and risks salient in the present or future" (p. 11). The final parenting practice Livingstone and Blum-Ross call *resist*. Here, parents try to stop media, screens, and technologies from becoming essential components of their children's lives, arguing that these present a problem for their children. National and international organizations have developed position statements for parents (e.g., the American Academy of Pediatrics [2016], Zero to Three [2018]) and teachers (e.g., the National Association for the Education of Young Children and the Fred Rogers Center [2012], the International Literacy Association [2018, 2019], and the National Council of Teachers of English [2019]) as well as curriculum programs (e.g., Common Sense Media's *Digital Citizenship Curriculum* [2020]) that present research findings and are designed to teach about staying safe online.

Several of these resources directly address privacy and security, offering guidance around practices and strategies to help children learn about keeping their information private and secure. What is key in each of these is that the adult holds the ultimate decision-making power in determining how, whether, why, and when a child may engage with digital technologies and media. None, however, present children with information about the digital environments and allow for informed children to make choices that reflect their critical understanding of the issues. Readers of this chapter may question whether children can understand, for example, how their worldviews can be limited by geofencing and other algorithmic tools that are driven by for-profit purposes. We decided to test the waters with our own nine children, who ranged in age from 4 to 12 years at the time of data collection. Cognizant of the parenting strategies of embrace, balance, resist, we viewed each parent-child dyad as a case study and asked the following questions: (1) How can parents and children understand and navigate the trends, forces, and tensions around privacy, security, and algorithms in their lives? and (2) In what ways might children become more reflective about the activities in which they engage on screens? Here we focus on our middle school children, who at the time of this data collection were approaching 13 years old, a critical age as defined by the Children's Online Privacy Protection Act. Table 3.1 displays the one parent-child dyad from each researcher's home that was included for this multiple case study.

By examining critical moments from each selected case, we were able to better understand how parents might engage middle school–aged children in conversations to better understand their view of digital technologies. These moments came from four dyads that included a researcher-parent and a single child (see table 3.1).

Our parent-child dyads took a range of approaches to generating conversations and data. These included guided drawing, graphic organizers, close reading and discussion of terms and services, and mentor texts around digital media and its use. We watched each video recorded and transcribed our interactions.

To analyze the data, we leveraged grounded theory and an open-coding approach (Corbin & Strauss, 2014). First, each of us open-coded our own transcripts. Then we shared and exchanged transcripts, meeting regularly to discuss the data. Our discussions centered on identifying the approaches that were effective in

Table 3.1
Parent-child dyads

Researcher/parent	Child	Age of child	Context of critical moment
Kristen	Megan	12	Reviewing the terms of use of an app that the child requested to download
Kathleen	Charlie	11	Discussion of privacy in social media apps in response to emails from concerned school administrators
Ian	Jax	9	Addressing the challenges of interactions with strangers when the child received a message online
Elizabeth	Addy	11	Reviewing risks and rewards of internet use in response to child's request for a smartphone

eliciting discussion and critical reflections in our children, specifically about understanding and navigating the trends, forces, and tensions around privacy, security, and algorithms in our lives.

The collaborative discussions provided a space for us to build consensus across the four case studies (Yin, 2017). Finally, we each recorded a reflective discussion with our child in the form of a podcast to allow our children a voice in the research process. These conversations revealed their perspectives on what they learned as participants and allowed them the opportunity to check our own understanding of their experience. We published versions of these recordings, as well as researcher reflections, publicly on our website (https://screentime.me/digital-futures/) and shared them with our social media networks in order to solicit feedback.

Through a collaborative, inductive approach that drew from our dual roles as parents and literacy researchers, we identified critical moments that highlighted themes that appeared across the data. These themes identify strategies parents and children can use to understand and navigate issues of privacy, security, and algorithms in digital spaces. These strategies include finding an approach point, providing media mentorship, assessing concerns head-on, and using language that empowers.

Find an Approach Point

Megan (age 12) owned a smartphone but was not yet a social media user. She was vocal about the effect social media had on her friends, and she had no interest in joining the bandwagon. However, she surprised her mother, Kristen (Author), after school one day by asking, "Can I get Snapchat?" Based on family rules, it would have been easy for Kristen to restrict, answer, "No," and move on.

However, Megan's question provided the perfect *approach point* for Kristen to discuss the roles of privacy, security, and algorithms on social media platforms. An approach point is a time, condition, or opportunity for a teachable moment, often through conversation. They sat together, perusing the terms of use and privacy agreements on the Snapchat website, and as they read together, Kristen clarified unfamiliar terms and concepts.

For example, the pair discussed the data that Snapchat collected and how algorithms allowed the company to use those data "to serve ads you might be interested in—when you might be interested in them" (Snap, 2019). By considering Megan's question, rather than responding with restricting and an immediate answer, Kristen was able to engage her daughter in conversation that helped her to understand the role of algorithms in the app her friends were using.

Provide Media Mentorship

Identifying the approach point with children is an important first step in teaching them about technologies. During these conversations, parents can provide *media mentorship*, or a guide that can help youth navigate the digital world while working to translate these experiences into positive and productive lifelong learning skills (Haines, Campbell, & the Association for Library Services to Children [ALSC], 2016).

Kathleen (Author) adopted a practice of "think-aloud" with her son Charlie (age 11) to provide mentorship. She invited Charlie to help shop for a new hockey stick, an activity that is oftentimes done in brick-and-mortar but was shifted to the internet to invite Charlie to examine and think critically about how algorithms function. While looking at reviews on YouTube, Kathleen pointed out the ads appearing in the margin. She thought aloud as she invited Charlie to observe: "I notice these boxes here don't seem to be related to hockey. They show me things that are a lot like what I've been searching for lately—vacations, proper grammar explanations to share with students. I wonder what might happen to these ads if you keep looking for things that interest you?"

By engaging Charlie in a routine task—shopping online—Kathleen was able to share her own thought process as she encountered ads while simultaneously prompting Charlie to think about the underlying algorithms. Similar mentorship can be done using books, TV shows, movies, and games as parents and children create, do, and explore together in order to help children better understand the workings of the internet and how algorithms affect what they see.

Address Concerns Head-On

As conversations between parents and children evolve, it is likely that issues about "safe spaces" will emerge. Ian learned that addressing concerns directly through explanations of algorithms, privacy, and security helped turn potential fear into vigilance.

Jax (age 9) was using Google Hangouts, an instant messaging platform that allowed him to share messages, photos, and videos with his parents. Though the family thought the account was completely private, accessible only to Jax and his parents, Jax was surprised to see a message from a stranger asking for photos of the child. Worried, Jax asked his father, Ian (Author), "Daddy, who sent me this message? Is it someone from the games I play?"

Though Ian took steps to protect Jax by blocking the account, he also recognized the moment as an approach point and explained to his son (and his younger daughter, 4 years old) how this message may have appeared. Their conversation about privacy, security, and algorithms allowed the children to adopt a stance of vigilance. Ian applauded Jax for bringing the breach to his attention, and instead of simply protecting his child by blocking an account, he addressed the concern head-on, bringing awareness to his children.

Use Language That Empowers

Parental instinct is to protect their children from harm, and it is tempting to use language that presents the internet in dichotomies: good/bad, safe/unsafe. Elizabeth learned that the language she used to talk about issues of privacy and security with her daughters mattered.

After reading *A Smart Girl's Guide: Digital World* (Anton, 2017) together and discussing the issues it raised, Elizabeth asked Addy (age 11) to explain what she learned. Addy said: "Not to use your real name, a photo of you, or pictures of your life. You need to be specific with passwords and accounts so you can stay safe. Sometimes you click on things that are not safe."

In reviewing the transcript of their conversation, Elizabeth realized that her own language may have influenced Addy's learning that the internet is a place that may not be "safe" and that she may not have control over her encounters in unsafe spaces. Through reflection, Elizabeth understood that she needed to use language that empowered her children to be agents in internet use, as compared to passive participants who are controlled by technology.

Language of empowerment would position the internet as an environment within which children may harness tools to learn, to be entertained, and to associate with global discourse communities in online settings. It would suggest that individuals can grow and develop in positive ways when they learn about themselves and the world around them, and it would celebrate individuals' internet use and expertise. During the recent US election, for instance, Addy used the internet to search for and identify media that supported her learning about current-event politics. As a result, Elizabeth and Addy were able to celebrate what she learned and use her new knowledge as an approach point to talk about history, worldviews, and policy. Instead of positioning the internet as "bad" or "unsafe," Elizabeth's use of empowering language positioned the internet as an environment within which her daughter could identify tools for learning.

Make Conversations Ongoing

Our research with our children has taught us that conversations about privacy, security, and the nature of algorithms need to start early and be ongoing. Both Megan and Charlie were able to articulate insight they gained from such conversations—specifically that most people do not know how algorithms work, and that virtually no one (especially none of their friends) reads terms of use and privacy agreements. Even so, they acknowledged that even if people knew more, it probably would not change how they use the internet because websites and apps are such an embedded part of life. Much of this discussion is also a challenge for adults as they often do not pay attention to the responsibilities necessary as web-literate citizens. The focus of our inquiry, and lessons learned from this work, point to the need for active knowledge construction and reflective practice as adults consider opportunities to empower youth to make informed, critical decisions about digital spaces and tools.

By early adolescence, our children are internalizing acceptable internet use. Parents and teachers need to be part of the conversation with them that shapes their understanding of these concepts. Jax was able to explain his knowledge to his 4-year-old sister, suggesting that this work can involve older children mentoring their younger siblings or schoolmates. This approach ultimately requires that parents and teachers open lines of communication with children as they strive to collaboratively make sense of these new environments. A restrictive approach (Livingstone & Blum-Ross, 2020) might not allow spaces for such critical and collaborative sensemaking, and likewise, parents who take either an embracing or a balancing stance might consider the critical role of conversation and child empowerment in family decisions.

As literacy researchers, we are parents with, perhaps, more knowledge about how algorithms and privacy work in a digital world. We recognize that not all parents may feel qualified to function as media mentors—to be honest, we did not either. Yet we sat at an interesting intersection (Garcia et al., 2014) in which we did not entirely view ourselves as experts or mentors in digital texts and tools even though that is one of our main areas of research and education. To address this, we proposed a more collaborative approach to mentoring around children, media, and technology than what has typically been adopted. Rather than framing the problem as technology doing harm to children, we suggest that we can empower children to advocate for their own rights in an age of screen time (Turner et al., 2017). We propose that all parents are experts in their children. If embrace, trust, and curiosity to learn more about the child's media interests are centered in the mentorship, the four strategies we have identified can support this effort regardless of a parent's perceived level of expertise.

References

American Academy of Pediatrics Council on Communication and Media. (2016). Media and young minds. *Pediatrics*, *138*(5), 1–6. https://doi.org/10.1542/peds.2016-2591

Anton, C. (2017). *A smart girl's guide: Digital world*. American Girl.

Berson, I. R. (2006). Children and their digital dossiers: Lessons in privacy rights in the digital age. *International Journal of Social Education*, *21*(1), 135–147. https://www.learntechlib.org/p/60795/

Bertot, J. C., Jaeger, P. T., & Hansen, D. (2012). The impact of policies on government social media usage: Issues, challenges, and recommendations. *Government Information Quarterly*, *29*(1), 30–40. https://doi.org/10.1016/j.giq.2011.04.004

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951715622512

Common Sense Media. (2020). *Digital citizenship curriculum*. https://www.commonsense.org/education/digital-citizenship/curriculum

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

Dillon, R. S. (2010). Respect for persons, identity, and information technology. *Ethics and Information Technology*, *12*(1), 17–28. https://doi.org/10.1007/s10676-009-9188-8

Emejulu, A., & McGregor, C. (2019). Towards a radical digital citizenship in digital education. *Critical Studies in Education, 60*(1), 131–147. https://doi.org/10.1080/17508487.2016.1234494

Garcia, A., Cantrill, C., Filipiak, D., Hunt, B., Lee, C., Mirra, N., O'Donnell-Allen, C., & Peppler, K. (2014). *Teaching in the connected classroom*. Digital Media and Learning Research Hub. https://dmlhub.net/publications/teaching-connected-learning-classroom/index.html

Haines, C., Campbell, C., & Association for Library Services to Children. (2016). *Becoming a media mentor: A guide for working with children and families*. ALA Editions.

International Literacy Association. (2018). *Improving digital practices for literacy, learning, and justice: More than just tools* [Literacy leadership brief].

International Literacy Association. (2019). *Digital resources in early childhood literacy development* [Position statement and research brief].

Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., Schor, J., Sefton-Green, J., & Watkins, S. C. (2013). *Connected learning: An agenda for research and design*.

Digital Media and Learning Research Hub. https://dmlhub.net/publications/connected-learning-agenda-for-research-and-design/

Livingstone, S., & Blum-Ross, A. (2020). *Parenting for a digital future: How hopes and fears about technology shape children's lives*. Oxford University Press.

National Association for the Education of Young Children & Fred Rogers Center for Early Learning and Children's Media at Saint Vincent College. (2012). *Technology and interactive media as tools in early childhood programs serving children birth through age 8*. http://www.naeyc.org/files/naeyc/file/positions/PS_technology_WEB2.pdf

National Council of Teachers of English. (2019). *Definition of literacy in a digital age*. https://ncte.org/statement/nctes-definition-literacy-digital-age

Snap. (2019). *Snapchat privacy policy*. Retrieved February 25, 2020, from https://www.snap.com/en-US/privacy/privacy-policy

Turner, K. H., Jolls, T., Hagerman, M. S., O'Byrne, W. I., Hicks, T., Eisenstock, B., & Pytash, K. E. (2017). Developing digital and media literacies in children and adolescents. *Pediatrics*, *140*(Suppl. 2), S122–S126. https://doi.org/10.1542/peds.2016-1758P

Willson, M. (2018). Raising the ideal child? Algorithms, quantification and prediction. *Media, Culture & Society*, *41*(5), 620–636. https://doi.org/10.1177/0163443718798901

Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage Publications.

Zero to Three (2018, October 25). *Screen-use tips for parents of children under three*. https://www.zerotothree.org/resources/2531-screen-use-tips-for-parents-of-children-under-three#downloads

4

Parenting and the Algorithm

A Perspective on Parental Controls and Guilt amid Digital Media

Maureen Mauk

Parental guilt. As a former broadcast television and children's content standards and practices (S&P) censor and a parent of two young kids, it took entering academia for me to upgrade this term with what I now call *responsibilization*. When it comes to digital media, according to the social rules governing public scripts, not only must I provide my child with access and the skills to manage content, but I am also held responsible to protect them from it.

When I turn on a Disney+ program for my six-year-old to buy myself an hour's time to do work, I take note of the "outdated cultural depictions" and "contains tobacco use" disclaimers appearing on the app. As a former television censor, I think about the negotiations and decision-making that had to happen in order to label this content and provide this type of warning over classic content carrying antiquated and racist portrayals. As a parent in a pandemic, hoping to enable my child to pick their own programming but unable to hover in the room to better explain these depictions, I wish I could take these choices off the menu altogether. Finally, as a media researcher, I recognize that the norms of the platform and the affordances of the technology allow me, the parent, very little customizable utility.

When considering children's rights and the obligations corporations and regulators have to them in the algorithmic environment, we must not forget about the parents. Media regulation such as the Children's Online Privacy Protection Act (COPPA) and the Children's Television Act are traditionally framed as advocating *for the children* with the expressed intent to protect society's most "vulnerable subjects" (Sefton-Green 2006, p. 282). Yet the trope fails to consider the needs of parents playing the role of familial gatekeeper, where *good parents* are implicated as the primary wave of defense to protect children from modern media.

Parents carry limitations and hesitations on what they can control in the digital realm and how they might make digital platforms, apps, and parental controls work best within their household. Technology carries constraints as well, and when it comes to streaming and social media, there is no one-size-fits-all customization of features. But in a self-governed, data-rich environment, the platforms and industry providers possess the power to not only better protect children but also to ease the burden on parents. Parents need advocates in a broader regulatory arena to voice their concerns as they are often so busy handling their own kids' media they do not have time to push for a major course correction in digital rights for stakeholders.

Using my perspective as a media scholar, a former television S&P executive, and a parent, I point to some of the issues and implications parents face in carrying the responsibility for children's digital media. I specifically approach my argument from the angle of digital parental controls in the United States, which have evolved from network television program practices and self-regulation. I first discuss the positioning of good parents and the ideology behind their responsibilities. Then, I offer a vantage point from the parent perspective of evaluating the design and affordances of parental controls. Finally, I call for parents and caregivers' voices to be better amplified in the industry and regulatory arena.

Responsibilization and the Good Parent

Long before the 2020 pandemic, parents have navigated their media responsibilities as part of the *pull yourselves up by your bootstraps* American ideology. The success of our children has been governed through a political rationale that interpolates individual families as in charge of their destiny (Willett, 2015; Cowan, 1983; Pugh, 2009). These practices, however, are not new. Meredith Bak (2020) discusses the preoccupations by parents in the nineteenth century for using new media toys (then in the form of a zoetrope!) to educate youth. In Daniel Thomas Cook's book *The Moral Project of Childhood*, designing a malleable child through consumer culture in the nineteenth-century United States centered on "maternal responsibility," what he calls "the moral project" (2020, p. 4). Part of the contemporary role of the caregiver is to provide technological opportunities that will ensure a child's success—from preschool edutainment and private day care to coding classes and SAT test prep, and everything in between (Ito, 2012; Hoover et al., 2012; Livingstone & Blum-Ross, 2020). Mothers and fathers do not balk at this responsibility because it feels common sense, inherent to the job of being a good parent. But parents cannot simply *provide* the technology; they also must *protect* their children from it. Parents furnish and then must regulate the child's digital realm. The *responsibilization* of parenting is a term that describes not only the guilt but also the flood of advice and assumed duties that accompany our surge in digital technology. Responsibilization is an ideology of parenting protectionism that has saturated the tech and media industry, government regulation, and public discourse. In the world of digital parenting that we've come to accept as normal, the parent is the gatekeeper, the censor, and the protector of children amidst an onslaught of capitalism and self-regulation.

Within the media ecosystem, *good parents* have been constructed as part of the dominant discourse, heard throughout society to point out what a parent *should* be doing to monitor children's digital environments. Alicia Blum-Ross and Sonia Livingstone, whose extensive research delves into digital parenting practices in the UK, documented "confessions" of "laziness" and "sentiments of guilt" in the parents they sat down with, writing: "Time and again we heard parents of young children struggle to balance the convenience of screen time with their worries about being a 'good' parent" (2018, p. 183). Beyond television media, caregivers have also been held accountable in media discourse for overseeing their children's online activities, responsible for promoting educational and learning opportunities, which often favor middle- and upper-middle-class families and their media habits (Clark, 2013; Lareau, 2003). Media research has shown that the gender and classed hierarchies associated with *good parenting* fault parents if they aren't monitoring their children's shows and games, if they fail to set up parental controls, or if they overindulge their children in the consumer media marketplace (Clark, 2013; Seiter, 1995; Steiner & Bronstein, 2016; Willett, 2015).

Regulators, media providers, and even organizations set up to help protect and educate parents lean on the ideal of the *good parent*. The international nonprofit Family Online Safety Institute (FOSI), boasting members from across government and tech sectors ranging from Amazon to Verizon, created a free downloadable book entitled "How to Be a Good Digital Parent" for parents seeking guidance on technology in the home (2022). Our media and parenting culture has grown accustomed to the idea that a *good parent* is one who is righteously vigilant in watching over children's media consumption and digital experiences. The term *parental control* and its utility offered by many apps and platforms may offer families a sense of empowerment through a suite of technology affordances. But in reality, it is alleviating pressure from the digital provider, shifting industry self-regulation to the home by aiming squarely at those parents aspiring to be *good*.

Parental Controls, Power, and the Ineffective Toolkit

Rarely would a parent or caregiver describe the digital realms that our families operate in, particularly the parental controls offered by technology companies, as an oppressive constraint. If anything, our culture tells us it is a family's path to freedom, a choose-your-own-adventure. We have become so accustomed to the discussion and industry-created buzz surrounding digital affordances (e.g., parents have the "tools they need to make wise decisions about what is right for themselves and their families" [Netflix, 2018]) that it seems absurd to consider families oppressed. Yet whether parental controls are used or ignored, we must recognize that the "tools" offered represent the transfer of the regulatory burden from government to industry to the parent at home navigating kids' content.

When evaluating the tools offered to parents, we need to consider what is missing. Digital streaming services are quick to point out the offerings and personalization their platforms and upgrades provide parents. Netflix claims its algorithmic technology helps its members be "better informed, and more in control, of what they and their families choose to watch and enjoy on Netflix" (Hastings, 2018). Beyond the parent PIN code and baseline kids' profile maturity setting, the parental controls offer few actual controls to parents.

Given what we know about Netflix's use of algorithmic personalization based on user metadata (Seaver, 2018; Tryon, 2015) and its practice of tagging kids' content internally (Grothaus, 2018), the limitations of my control as a parent are just that—limited. I have no power to instruct a streaming platform to remove outdated cultural depictions, stories about fire or ghosts, or the phrase "shut up." As a *good parent*, I've bought into the belief that I need to do my due diligence to protect my children from various depictions and references. As a subscriber, I realize I have very little power to filter content, despite how powerful streamers have told me their algorithms and data might be. As a former censor studying media culture, I wonder how parents might handle these responsibilities while operating in the shadows of opaque offerings. If digital content providers are relegating self-regulation to parents, shouldn't parents be offered more tools to do so?

Advocating for Parents

There are too few safeguards or regulations surrounding platform governance to protect parents or ease their burden. The reality of digital parenting is one where kindergarten teachers assign videos via Seesaw platforms on iPads. The old adage to "just turn it off" won't cut it. Traditional swathes such as V-chip ratings also will not suffice. We live in a radically different media environment than we did in the early days (the nineties!) of the V-chip, where regulators impelled traditional television broadcasters to create standards and blocking functionality across linear TV programming. The digital environment and children's metaverse has enveloped the child-rearing experience. Its global but opaque nature has clouded more traditional pathways of protecting the end user through industry-wide regulation. Platforms and providers, however, need not wait for top-down regulation to better serve parents; they just need to pay better attention to how their data and affordances can best help families. And families need better advocates.

When I recently suggested at an international conference on social media governance that parents would have to "take to the streets" to push back and demand better offerings and services in the form of industry self-regulation and government guidelines, I was met with sympathetic chortles and snorts. It is laughable to imagine mothers taking to picket lines for this matter of contention amid the many issues we are all facing as global citizens. But that is my point. It is laughable, not because it is unimaginable but because the guilt that technology and parental controls have created can barely be a priority when families as a whole are not prioritized within the intensified and commercialized framework of US parenting in the digital age.

When thinking of the children, we also need to consider the parents and encourage research for civic-minded justice for families through the lens of domestic media practices. If parents are being guilted into manning the controls for kids' content, I argue that we must advocate for better controls. We should demand more of our lawmakers and the tech industry to marshal and cultivate data that supports personalized tools for parents and caregivers. To protect children, we must start by protecting parents. Advocacy for parents should reflect their wide and varied needs; it should become a focus across government, technology, and media sectors, working to promote increased transparency and accountability. To advocate for parents means we must recognize the limitations of technology and parental controls and work to lessen the guilt and burden of responsibility weighing on parents in the digital domestic arena.

References

Bak, M. (2020). *Playful visions*. The MIT Press.

Blum-Ross, A., & Livingstone, S. (2018). The trouble with "screen time" rules. In G. Mascheroni, C. Ponte, & A. Jorge (Eds.), *Digital parenting: The challenges for families in the digital age* (pp. 179–188). The International Clearinghouse on Children, Youth, and Media.

Clark, L. S. (2013). *The parent app: Understanding families in the digital age*. Oxford University Press.

Cook, D. T. (2020). *The moral project of childhood*. NYU Press.

Cowan, R. S. (1983). *More work for mother: The ironies of household technology from the open hearth to the microwave*. Basic Books.

Family Online Safety Institute. (2022). *Good digital parenting*. https://www.fosi.org/good-digital-parenting

Grothaus, M. (2018, March 28). *How I got my dream job of getting paid to watch Netflix*. Fast Company. https://www.fastcompany.com/40547557/how-i-got-my-dream-job-of-getting-paid-to-watch-netflix

Hastings, R. (2018, March 5). *Introducing PIN protection and other enhancements for informed viewing*. Netflix Media Center. https://about.netflix.com/en/news/introducing-pin-protection-and-other-enhancements-for-informed-viewing

Hoover, S. M., Clark, L. S., & Alters, D. M. (2012). *Media, home and family*. Routledge.

Ito, M. (2012). *Engineering play: A cultural history of children's software*. The MIT Press.

Lareau, A. (2003). *Unequal childhoods: Class, race, and family life*. University of California Press.

Livingstone, S., & Blum-Ross, A. (2020). *Parenting for a digital future*. Oxford University Press.

Netflix. (2018). *Maturity ratings for TV shows and movies on Netflix*. https://help.netflix.com/en/node/2064

Pugh, A. (2009). *Longing and belonging: Parents, children, and consumer culture*. University of California Press.

Seaver, N. (2018). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, *24* (4), 421–436. https://doi.org/10.1177/1359183518820366

Sefton-Green, J. (2006). Youth, technology, and media cultures. [Special issue: Rethinking learning: What counts as learning and what learning counts]. *Review of Research in Education*, *30*, 279–306. https://www.jstor.org/stable/4129775

Seiter, E. (1995). *Sold separately: Children and parents in consumer culture*. Rutgers University Press.

Steiner, L., & Bronstein, C. (2016). When tiger mothers transgress: Amy Chua, Dara-Lynn Weiss and the cultural imperative of intensive mothering. In H. L. Hundley & S. E. Hayden (Eds.), *Mediated moms: Contemporary challenges to the motherhood myth* (pp. 247–273). Peter Lang Publishing.

Tryon, C. (2015). TV got better: Netflix's original programming strategies and binge viewing. *Media Industries*, *2*(5), 104–116. https://doi.org/10.3998/mij.15031809.0002.206

Willett, R. J. (2015). The discursive construction of "good parenting" in the case of children's virtual world games. *Media, Culture & Society*, *37*(7), 1060–1075. https://doi.org/10.1177/0163443715591666

5

Meet Them Where They Are

Social Work Informed Considerations for Youth Inclusion in AI Violence Prevention Systems

Desmond U. Patton, Siva Mathiyazhagan, and Aviv Y. Landau

Introduction

Artificial intelligence (AI) systems are being created to monitor and predict youth violence that occurs globally on social media. Some youth use social media as a psychosocial tool for help-seeking, grief processing, and general support. Conversely, other youth may use social media to exclude others or engage in cyberbullying, violence, and additional acts of isolation. To combat these challenges, school districts, law enforcement, and criminal justice organizations are leveraging artificial intelligence (AI) (e.g., machine learning and computer vision) to identify and predict harmful content online (Patel et al., 2020).

While the concept of predicting and preventing harmful content online seems hopeful, there are deep concerns regarding the extent to which an AI system can correctly decipher context, which is critical to any interpretation of language or action. Due to the lack of understanding of language, cultural nuances, and social context, there are numerous impacts when AI technologies wrongly interpret youths' posts as violent. As such, AI systems may create and reinforce new systems of marginalization and oppression and even put youth at risk by creating digital pathways to incarceration (Patton et al., 2017). As social work researchers from the US, India, and Israel serving youth and researching AI in three different countries, we have a front-row seat to the transnational interactions between youth, technology, government, and the private sector. Our social work and lived experiences in these countries brought us together to discuss social work approaches in data science to prevent online violence against young people in our respective countries. We suggest that social work ethical principles of respect for diversity, human rights, anti-oppression, privacy, and safety be integrated broadly as a framing guide for developing AI technologies to prevent violence against and among youth and marginalized groups.

Technology and Youth Culture in India, Israel, and the United States

India

India has 600 million youth under 25 years old (Jack, 2018). There are around 645 distinct tribes in India and more than 19,500 languages and dialects spoken as mother tongues. While only 22 languages are officially listed in the Indian constitution, Google recognizes and supports only nine of these languages (Office of the Registrar General & Census Commissioner, India, 2011; Shukla, 2019). India is home to 560 million internet subscribers with 351 million active monthly social media users, which is predicted to double by 2023 (Pragati, 2019; PTI, 2019). Growing access to new age technology and free access to social media have opened new spaces for young people to express their feelings and thoughts on social media. As they engage with these digital platforms, artificial intelligence has been applied in various ways to Indian youths' daily life through social media monitoring, linking of biometric IDs with services, and digital marketing irrespective of caste, gender, sexuality, and religion (Chawla, 2020; Jalan, 2020; Singh & IANS, 2020). Nevertheless, Indian scholars worry that the new AI systems might reinforce caste and religious discrimination through modern tech bias in employment, imprisonment, and access to finances, similar to the consequences of racial bias in the US (Kalyanakrishnan et al., 2017).

Social activists are continuously concerned about insufficient regulations applied to emerging technologies (Ulmer & Siddiqui, 2020). An example of this is the Aadhaar ID, a digital biometric ID system that collects personal details like photos, fingerprints, and demographic profiles and links them with the individual's welfare and banking services (Jain, 2019; Pandya & Cognitive World, 2019). When linked with AI programs, Aadhaar ID allows the government to scan and flag certain citizens as suspicious or dangerous. State governments, like Punjab and Delhi, having already installed the Automated Facial Recognition System (AFRS) software in airports, offices, and cafes to identify "criminals," are now extending AI-enabled facial recognition algorithms to screen crowds at political rallies and people's protests (Chandran, 2019; Jain, 2019; Ulmer & Siddiqui, 2020). During recent protests against citizenship law, young people covered their faces because they were afraid police were using facial recognition systems to identify and arrest them. The Indian National Strategy for Artificial Intelligence discussion paper acknowledges that data-driven decision-making and AI algorithms in the country may be biased, and it is important to critically assess the impact of these biases on society and find ways to reduce them (NITI Aayog, 2021). Without addressing built-in bias, there is a significant chance that AI may mispredict youth expressions in local or regional languages on social media. Moreover, because there is no policy in India to ensure safe, inclusive, and participative AI technologies for young people, the rise of AI in India might cause offline violence and increase polarization in society. Widespread rumors on social media have already aggravated many acts of communal violence, including the Dadri Mob Lynching in 2015 and the Kathua Rape Case in 2018 (Teitelman, 2019). AI needs caste-, gender-, and community-centered contextualization in Indian sociocultural context to prevent possible bias and harm.

Israel

During the past decade, social media has changed how Israeli youth socialize. Although youth have more possibilities to enhance existing friendships and engage in new relationships, they still experience exclusion, cyberbullying, and other violent behaviors on social media. The unique makeup of Israeli society—comprising different beliefs, cultures, and norms—may further exacerbate these on- and offline tensions (Aizenkot & Kashy-Rosenbaum, 2019; Landau et al., 2019; Mesch, 2017). Yet deploying AI technologies for safe online environments for youth is in its infancy. New Israeli high-tech companies, like L1ght, are developing AI systems for monitoring social media and the internet in the hope of preventing and protecting youth from cyberbullying, shaming, and sexual predation (Chaimovich, 2020). At the same time, the central law enforcement agency has plans to develop an AI system to monitor negative social media posts, such as threats, incitements, and online shaming, that are directed toward police officers. This plan is concerning, as it could provide the police with unlimited access to Israelis' social media without any restraint or consent. Israel currently lacks ethical AI policies and guidelines, igniting growing concern that companies and police have access to unrestricted surveillance, thus violating fundamental human rights of privacy and consent (Kabir, 2019). Without training AI technology to understand the different cultures, norms, and beliefs of youth within the country, the development of these technologies may reinforce biased assumptions that can lead to further exclusion and violence.

United States

In the United States, integrating and deploying AI technologies has fundamentally changed our lives and, in particular, the lives of young people. There has been tremendous discussion about the use and misuse of facial recognition systems, particularly for communities of color and transgender individuals. Much of this work has come into focus because of the research and advocacy of Black women, like Joy Buolamwini, Timnit Gebru, and Mutale Nkonde, who discuss the large racial bias in algorithmic systems that extend and amplify racial inequity (Buolamwini et al., 2020; Nkonde, 2019; Raji et al., 2020). These results of faulty facial recognition systems were underscored in a recent *New York Times* article that described the experience of a Black man who was falsely arrested in Detroit for a crime he did not commit (Hill, 2020). In addition to facial recognition, new research from the Brennan Center at NYU indicates that over the last five years, new surveillance companies have developed and are selling software, powered by AI, that can allegedly detect signs of violence or other concerning behavior among youth on social media (Patel et al., 2020). One example of implementing this technology is Chicago Public Schools. Armed with a US Department of Justice grant, the large urban district hired intelligence analysts and purchased a social media monitoring service to analyze online conversations among students. The analysts used keyword searches to find threats at the program's target schools (Patel et al., 2020). The program is particularly concerning because students were not made aware of this initiative and it remains unclear what words or phrases connote a "threat." This is precarious, given recent research from Patton and colleagues (2019) who found that Chicago youth from a neighborhood with high rates of gun violence did not agree on how to interpret Twitter posts identified as "aggressive" with peers from the same neighborhood. Nevertheless, research from the SAFElab documents back-and-forth arguments on social media between youth that, in some cases, lead to online aggression, school fights, increased bullying, and gun violence. Prevention is critical, and the AI violence prevention system needs local youth participation to find meaning, language, context, and communication styles, as much of the language and images used on social media are susceptible to misinterpretation. While it is exciting to leverage AI systems to identify potential harms, there is little to no evidence that suggests AI meets the goals for which it has been deployed.

Integration of Social Work Ethics in Technological Inclusion

Around the world, social workers are practicing a similar code of ethics, adopted by the International Federation of Social Workers (IFSW) and national associations, irrespective of the sociocultural complexities in the world. In India, recognizing and incorporating indigenous knowledge is essential in understanding youth and the complex local language expressions on social media and the internet. It is critical that AI systems are optimized to identify the pragmatic ways in which youth use social media and contextually understand languages, particularly from marginalized communities with myriad languages and hyperlocal context. Although there are no standard national social work ethical principles in India, the IFSW Code of Ethics is widely adopted and practiced across the country to ensure professional ethics in human services. The development, integration, and application of AI systems in India should prioritize principles that underscore human rights, social justice, community participation, equity, and ethical use of technology as highlighted in the IFSW Code of Ethics (IFSW, n.d.). These ethical considerations outline technology's role in social work practice, as well as offer scope for social work's role in developing and deploying emerging technologies in a real-time practice to foster inclusion and prevent violence.

Israel's Social Work Code of Ethics guides social workers in practice to support their clients', families', and communities' participation and quality of life (Israeli Association of Social Work [ISASW], 2018). Youth in Israel come from different beliefs, languages, races, and ethnicities, such as Jewish, Christian, and Muslim, Israeli-Arabs, and immigrants from different countries such as Ethiopia and the former USSR, and it is essential to obtain their insights around violence to reduce bias assumptions. Because AI technology for violence prevention has ethical connotations, developing and implementing AI technologies for youth violence prevention in Israel should consider adopting a social work ethical approach that involves youth participation from different sectors of the country to increase objectivity and ultimately develop more effective AI systems.

In the US, social workers follow the National Association of Social Workers (NASW) Code of Ethics, which frames everyday professional conduct and practice for social workers. At its core, the framework espouses ideas of service, social justice, personal dignity and worth, importance of human relationships, integrity, and competence. In 2018 the NASW, along with the Association of Social Work Boards, the Council on Social Work Education, and the Clinical Social Work Association, developed standards to consider technology's role in social work practice and education. The standards cover four main areas: provide information for the public, design and deliver services, gather and manage information about a client, and educate and supervise students. At its core, the relatively new standard is grounded in ethics, the text proclaiming, "When social workers use technology to provide information to the public, they shall take reasonable steps to ensure that the information is accurate, respectful, and consistent with the NASW Code of Ethics (NASW Cultural Standards, pp. 16)" (NASW, 2017). Let's take for example the use of AI for predictive risk assessment in child welfare. Social work researchers from the Children's Data Network, a research initiative at the University of Southern California's Suzanne Dworak-Peck School of Social Work, have used AI to link health records across all aspects of a child's life, including health, education, or Department of Child and Family Service (DCFS) data, with the goal of improving the well-being of children—perhaps better identifying children at risk before serious injury or death—and influencing policy decisions (Cuccaro-Alamin et al., 2017). While there are many ethical considerations to contend with, a social work approach might consider working with directly impacted groups as domain experts to co-design those AI systems. This means seeking qualitative insights and expertise to consider and anticipate potential challenges, harms, or benefits of an AI deployment. It is also critical to consider how issues of privilege, oppression, race, and power play out in creating data sets; how and what labels or codes are created; and how and where the AI system is deployed.

AI systems are used in India, the US, and Israel without, or with limited, sociocultural contexts. Incorporating social work ethical principles and an anti-oppressive approach can add sociocultural contextual value in AI development and integration, as shown in table 5.1. These ethical considerations can outline technology's role in social work practice as well as offer scope for social workers' role in developing and deploying emerging technologies beyond borders to promote inclusion and prevent violence.

Conclusion

AI tools are used in India, Israel, and the United States under the guise of youth violence prevention. While there is evidence that problematic content does occur on social media across populations and platforms, there is a dearth of evidence that suggests AI can actually reduce harmful and hateful content online. The field

Table 5.1
Transnational approach to social work ethics in emerging technologies

Transnational approach	India	US	Israel
Tech challenges in society	No policy to regulate AI technologies, which leads to potential surveillance, digital victimization, e-incarceration, bias and discrimination, and the digital divide.	No federal policy to regulate AI technologies, which leads to potential surveillance, digital victimization, e-incarceration, bias and discrimination, and the digital divide.	Israel lacks ethical AI policies and guidelines, potentially providing tech companies and police access to unrestricted surveillance.
Key social differences	Caste is an invisible systemic social problem. Religious polarizations, diverse languages, and cultural groups.	Race is the most visible structural systemic social problem. Multinational cultural groups and indigenous populations.	Systemic issue between Jewish and Arab ethnic and religious populations.
Social workers' engagement in the welfare systems	Social workers are actively engaged in the welfare systems.	Social workers are actively engaged in the welfare systems.	Social workers are actively engaged in the social welfare systems.
Social work ethics	India adopts ethical principles from the International Federation of Social Workers (IFSW) that highlights human rights values, social justice, anti-oppressiveness, people's participation, self-determination, diversity, and indigenous knowledge (IFSW, n.d.).	NASW Code of Ethics highlights the role of technology in social work practice and education, emphasizing human rights values, social justice, anti-oppressiveness, people's participation, self-determination, diversity, and indigenous knowledge (NASW, 2017).	The ISASW guides highlight social work practice that emphasizes human rights values, social justice, anti-oppressiveness, people's participation, self-determination, and diversity (ISASW, 2018).
Potential transnational benefits of tech–social work collaboration	• Adopting social work ethical principles in tech development, tech deployment, and use of data to reduce harm and bias. • Bringing diverse voices into the emerging technology. • Promoting community participation. • Building safe and inclusive technologies with social workers.		

of social work offers prevention and intervention models that provide a framework and context for working with diverse populations, particularly through leveraging domain expertise, practicing tech social work, and promoting social cohesion. Shared globally, these varied social work principles may offer a more ethical and humane approach to developing AI technologies and tools for violence prevention. They may also serve as a check against using AI when the tool does not fit the social problem, the research question, or the social context. With appropriate ethical standards in place, social workers, computer scientists, local youth, and other stakeholders of youth development can co-create and collaborate to prevent harm and bias in the AI systems.

References

Aizenkot, D., & Kashy-Rosenbaum, G. (2019). Cyberbullying victimization in WhatsApp classmate groups among Israeli elementary, middle, and high school students. *Journal of Interpersonal Violence*, *36*(15–16), NP8498-NP8519. https://doi.org/10.1177/0886260519842860

Buolamwini, J., Ordóñez, V., Morgenstern, J., & Learned-Miller, E. (2020). *Facial recognition technologies: A primer*. Algorithmic Justice League. https://global-uploads.webflow.com/5e027ca188c99e3515b404b7/5ed1002058516c11edc66a14_FRTsPrimerMay2020.pdf

Chaimovich, H. (2020, February 25). *The "shark" that will protect children: Zohar Lebovitz's L1ght raised 15 million dollars*. Geek Time. https://www.geektime.co.il/l1ght-raised-15m/

Chandran, R. (2019, December 30). *Use of facial recognition in Delhi rally sparks privacy fears*. Reuters. https://www.reuters.com/article/us-india-protests-facialrecognition-trfn/use-of-facial-recognition-in-delhi-rally-sparks-privacy-fears-idUSKBN1YY0PA

Chawla, V. (2020). Is social media analytics redundant today? *Analytics India Magazine*. https://analyticsindiamag.com/is-social-media-analytics-redundant-today/

Cuccaro-Alamin, S., Foust, R., Vaithianathan, R., & Putnam-Hornstein, E. (2017). Risk assessment and decision making in child protective services: Predictive risk modeling in context. *Children and Youth Services Review*, *79*(August), 291–298. https://doi.org/10.1016/j.childyouth.2017.06.027

Hill, K. (2020, June 24). Wrongfully accused by an algorithm. *The New York Times*. https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html

International Federation of Social Work (IFSW). (n.d.). *Statement of ethical principles and professional integrity*. https://www.ifsw.org/wp-content/uploads/2018/06/13-Ethics-Commission-Consultation-Document-1.pdf

Israeli Association of Social Work (ISASW). (2018). *The Israeli social work code of ethics*. https://socialwork.org.il/prdFiles/%D7%A7%D7%95%D7%93%20%D7%94%D7%90%D7%AA%D7%99%D7%A7%D7%94%20%D7%94%D7%97%D7%93%D7%A9%202018.pdf

Jack, I. (2018, January 13). India has 600 million young people—and they're set to change our world. *The Guardian*. https://www.theguardian.com/commentisfree/2018/jan/13/india-600-million-young-people-world-cities-internet

Jain, M. (2019, May 19). *The Aadhaar card: Cybersecurity issues with India's biometric experiment*. The Henry M. Jackson School of International Studies, University of Washington. https://jsis.washington.edu/news/the-aadhaar-card-cybersecurity-issues-with-indias-biometric-experiment/

Jalan, T. (2020, June 11). *Indian government again proposes social media surveillance, this time in the name of fake news*. MediaNama. https://www.medianama.com/2020/06/223-india-social-media-surveillance-proposal-fake-news/

Kabir, O. (2019, October 7). *Police shaming procedure: Monitoring and removing citizens' statements against police*. Calcalist. https://www.calcalist.co.il/internet/articles/0,7340,L-3771613,00.html

Kalyanakrishnan, S., Panicker, R. A., Natarajan, S., & Rao, S. (2017). Opportunities and challenges for artificial intelligence in India. *AIES Conference 2018*. https://www.aies-conference.com/2018/contents/papers/main/AIES_2018_paper_52.pdf

Landau, A. Y., Eisikovits, Z., & Rafaeli, S. (2019). Coping strategies for youth suffering from online interpersonal rejection. *Hawaii International Conference on System Sciences*, *52*, 2176–2185. https://scholarspace.manoa.hawaii.edu/items/800d935e-d3de-4955-a7e4-f53e6e39d4fe/full

Mesch, G. S. (2017). Race, ethnicity and the strength of Facebook ties. *Journal of Youth Studies*, *21*(5), 575–589. https://doi.org/10.1080/13676261.2017.1396303

National Association of Social Work (NASW). (2017). *The NASW Code of Ethics*. https://www.socialworkers.org/About/Ethics/Code-of-Ethics/Code-of-Ethics-English

NITI Aayog. (2021). *National strategy for artificial intelligence #AIforall*. https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf

Nkonde, M. (2019). Automated anti-Blackness: Facial recognition in Brooklyn, New York. *Harvard Kennedy School Journal of African American Policy*, *20*, 30–36. https://hjaap.hkspublications.org/wp-content/uploads/sites/14/2020/05/Final-PDF-for-Website.pdf

Office of the Registrar General & Census Commissioner, India. (2011). *2011 census data [data set]*. Ministry of Home Affairs, Government of India. https://censusindia.gov.in/census.website/data/census-tables

Pandya, J., & Cognitive World. (2019). Nuances of Aadhaar: India's digital identity, identification system and ID. *Forbes*. https://www.forbes.com/sites/cognitiveworld/2019/07/16/nuances-of-aadhaar-indias-digital-identity-identification-system-and-id/?sh=85a2ed1209da

Patel, F., Levinson-Waldman, R., Koreh, R., & DenUyl, S. (2020). *Social media monitoring: How the Department of Homeland Security uses digital data in the name of national security*. Brennan Center for Justice. https://www.brennancenter.org/our-work/research-reports/social-media-monitoring

Patton, D. U., Blandfort, P., Frey, W. R., Gaskell, M. B., & Karaman, S. (2019). Annotating Twitter data from vulnerable populations: Evaluating disagreement between domain experts and graduate student annotators. *Hawaii International Conference on System Sciences*, *52*, 2142–2151. https://www.researchgate.net/publication/330261697_Annotating_Twitter_Data_from_Vulnerable_Populations_Evaluating_Disagreement_Between_Domain_Experts_and_Graduate_Student_Annotators

Patton, D. U., Brunton, D., Dixon, A., Miller, R. J., Leonard, P., & Hackman, R. (2017). Stop and frisk online: Theorizing everyday racism in digital policing in the use of social media for identification of criminal conduct and associations. *Social Media + Society*, *(3)*3. https://doi.org/10.1177%2F2056305117733344

Pragati. (2019). *Social media statistics in India*. Talkwalker. https://www.talkwalker.com/blog/social-media-statistics-in-india

PTI. (2019, April 25). Internet users in India to rise by 40%, smartphones to double by 2023: McKinsey. *The Economic Times*. https://economictimes.indiatimes.com/tech/internet/internet-users-in-india-to-rise-by-40-smartphones-to-double-by-2023-mckinsey/articleshow/69040395.cms

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J, & Denton, E. (2020). Saving face: Investigating the ethical concerns of facial recognition auditing. *AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 145–151. https://doi.org/10.1145/3375627.3375820

Shukla, G. (2019, September 20). *Google search to add support for 3 new Indian languages by end of this year, updated mobile search UI also coming*. NDTV: Gadgets 360. https://gadgets.ndtv.com/internet/news/google-search-oriya-urdu-support-3-new-languages-station-wi-fi-gram-panchayat-2104303

Singh, S. K., & IANS (2020). Govt exploring use of AI to tackle social media misuse. *Express Computer*. https://www.expresscomputer.in/artificial-intelligence-ai/govt-exploring-use-of-ai-to-tackle-social-media-misuse/45570/

Teitelman, C. (2019, April 19). Communal violence, social media, and elections in India. *Columbia | SIPA: The Journal of International Affairs*. https://jia.sipa.columbia.edu/online-articles/communal-violence-social-media-and-elections-india

Ulmer, A., & Siddiqui, Z. (2020, February 17). *India's use of facial recognition tech during protests causes stir*. Reuters. https://www.reuters.com/article/india-citizenship-protests-technology-idINKBN20B0ZK

Part II Full Papers

6
Designing for Critical Algorithmic Literacies

Sayamindu Dasgupta and Benjamin Mako Hill

Introduction

As pervasive data collection and powerful algorithms increasingly shape children's experiences, children's ability to interrogate computational algorithms is increasingly important. A growing body of work has sought to identify and equip children with the intellectual tools they might use to understand, interrogate, and critique powerful algorithmic systems. We call the intellectual tools that allow children to understand and critique these systems that affect their lives *critical algorithmic literacies*. Unfortunately, because many powerful algorithms are invisible, developing these literacies remains a major challenge. However, it is possible for designers to build systems to support the development of critical algorithmic literacies in children.

Reflecting on extensive observation and design work in the Scratch online community over the last decade, we offer four principles for designers that describe ways to support children in developing critical algorithmic literacies:

1. Enable connections to data
2. Create sandboxes for dangerous ideas

3. Adopt community-centered approaches
4. Support thick authenticity

Our first principle encourages designers to *enable connections to data* by offering children opportunities to engage directly in data analysis, especially with data sets that relate to the world the children live, learn, and play in. The rationale for this principle is that in an increasingly data-driven world, understanding algorithms is deeply connected to understanding data. As children analyze data in order to ask and answer their own questions or pursue their own interests, they create their own algorithms. Through this process, they can start to interrogate both their data and their algorithms.

Our second principle suggests that the development of critical algorithmic literacies can be supported by *creating sandboxes for dangerous ideas*. Algorithms are both powerful and potentially problematic. Our design work suggests that children can develop a deep understanding of both facts when they are allowed to create and experiment with algorithms using carefully designed toolkits. Because these toolkits allow learners to "play with fire" in ways that might lead to negative outcomes, effective toolkit design needs to ensure that the possible dangers are managed and minimized. We use the metaphor of "sandboxes" to describe the goal of managing risk in this design process.

Our third principle suggests that designers should *adopt community-centered approaches* that allow designs to leverage community values that algorithms might challenge. Children belong to many overlapping communities and typically share many of their communities' values. Algorithms are seen as problematic, by children and by society in general, when they violate these socially constituted values. A community-centered approach intentionally situates algorithms within communities that have particular sets of shared values. Doing so makes the problematic nature of algorithms visible to learners who are likely to be aligned with community values that an algorithm violates or challenges.

Finally, we argue that *supporting thick authenticity*—a principle that applies to learning technology design in general—plays a crucial role in developing critical algorithmic literacies. Authenticity in the context of fostering algorithmic or data literacies might mean engaging in activities that involve "real-world" data or scenarios.

First, we describe the theoretical work that informs the way we conceptualize critical algorithmic literacies as well as the empirical and design work we have conducted that has informed our design principles. Next, we describe and situate the four design principles with detailed examples. Finally, we discuss our principles' implications for future design work and conclude with a reflection on unanswered questions and future directions.

Background

Our work draws from the literature on constructionism, a framework for learning and teaching that emphasizes contexts of learning "where the learner is consciously engaged in constructing a public entity, whether it's a sand castle on the beach or a theory of the universe" (Kafai, 2006; Papert & Harel, 1991, p. 1). We are particularly inspired by Resnick and Silverman (2005), who provide a series of design principles for designing constructionist learning environments and toolkits based on reflections on their practice as designers. We have attempted to follow in Resnick and Silverman's footsteps by laying out design principles for critical algorithmic literacies.

We use the term *algorithmic literacies* to describe a subset of computational literacies as articulated by diSessa (2001) in his book *Changing Minds: Computers, Learning, and Literacy*. diSessa suggests three broad pillars for literacy—material, mental or cognitive, and social. Material involves signs, representations, and so on. For language literacy, the material pillar might include alphabets, syntax, and writing conventions. For computational literacies, the material might involve user interface paradigms like spreadsheets or game genres, or modes of transmission like sharing on social media. The second pillar—mental or cognitive—represents the “coupling” (p. 8) of the material and what goes on inside learners’ minds when interacting with the material. The final pillar—social—represents communities that form the basis of literacies. diSessa posits that the emergence of a given literacy is driven by “complex social forces of innovation, adoption, and interdependence” (p. 11).

More recently, Kafai et al. (2019) have proposed a framework with three frames for understanding computational thinking: the cognitive, the situated, and the critical. They call for approaches to computational thinking that integrate “cognitive understanding” in three forms: comprehension of computational concepts; “situated use,” meaning that learning happens in contexts the learner cares about; and “critical engagement” to emphasize why we must question the larger structures and processes behind the phenomenon being analyzed. These three frames can also be used in the context of computational literacies. In fact, one of the case studies used by Kafai et al. to illustrate their framework is framed around the concept of “critical data literacies” drawn from our work (Hautea et al., 2017).

Our use of the term “critical” draws from Agre’s (2014) idea of “critical technical practice,” which ties critique and questioning to the practice of building and creation. In that sense, our goal is not merely knowledge about algorithms (i.e., what algorithms are) but an ability to critique algorithmic systems reflexively. Agre posits critical technical practice as requiring a “split identity—one foot planted in the craft work of design and the other foot planted in the reflexive work of critique” (p. 155). We recognize that as children engage with our toolkits, their design work combined with their reflection allows them to not only understand technical concepts around algorithms (what Agre describes as “esoteric terms”) but also evaluate their implications on society (“exoteric terms”).

Finally, the notion of critical algorithmic literacies is rooted in Freire's (1986) literacy methods. As we use it, the term was first proposed by Tygel and Kirsch (2016), who noted parallels between Freirean approaches to literacy education and the potential of models for developing data literacy. In suggesting approaches to big data literacy, D'Ignazio and Bhargava (2015) also build on Freire to posit that "[big data] literacy is not just about the acquisition of technical skills but the emancipation achieved through the literacy process" (p. 5). Relatedly, C. H. Lee and Soep (2016) have described their extensive body of work with child-driven multimedia production at the "intersection of engineering and computational thinking on the one hand, and narrative production and critical pedagogy on the other" (p. 481) in terms of *critical computational literacy*. This is a framework developed by C. H. Lee and Garcia (2015) while studying children from south Los Angeles who created animations and interactive games about sociopolitical issues in their community, such as racial profiling.

Our design principles are the result of design and empirical research around two systems we have developed and deployed over the last 10 years: *Scratch Cloud Variables* and *Scratch Community Blocks*. Both tools were designed with constructionist framings of learning in mind. Both sought to support children in learning about computational concepts related to data collection, processing, and analysis. Both tools also built on and extended the Scratch programming language—a widely used, block-based programming language for children (Resnick et al., 2009)—and both were deployed in the Scratch online community, where Scratch users share, comment on, and remix their Scratch projects (Monroy-Hernández & Resnick, 2008).

The primary design goal of Scratch Cloud Variables was to allow children to collect, record, and analyze data within Scratch (Dasgupta, 2013a). The primary goal of Scratch Community Blocks was to allow children to analyze their own social data directly (Dasgupta & Hill, 2017). Scratch Community Blocks enabled this goal by allowing Scratch users to access and analyze data from the Scratch online community website's database. In deploying both systems, we found that granting children programmatic access to data led them to not only learn the techniques of data analysis but also question and critique data-driven algorithms.

The empirical data that we draw from are from field deployment–based studies we conducted with members of the Scratch online community as well as from face-to-face workshops that we ran in the greater Boston area. For Scratch Cloud Variables, the deployment was part of a larger beta test of the Scratch 2.0 software. For Scratch Community Blocks, 2,500 beta testers were randomly selected from a pool of active Scratch users. Our studies involved observing Scratch projects and comments on projects, as well as seeking feedback through forum posts, surveys, and interviews. To help situate our findings, it is worth noting that the median age of Scratch users is 12 years old, and most are between 11 and 15. Although the distribution varies over time, around two-thirds of Scratch users describe their gender as male, and a small number (approximately 5 percent) do not report gender or self-report using nonbinary genders. Our sample of 2,500 participants in the Scratch Community Blocks was roughly gender balanced but similar in age to the general population of Scratch users (Dasgupta & Hill, 2017, p. 3625).

Design Principles

Over the last decade, much of our research has focused on designing, deploying, and studying systems that seek to support constructionist learning around data and data-intensive algorithms in Scratch. We distill lessons from this work into four principles that we believe will be useful for a range of designers interested in helping children learn about data-driven computational techniques, as well as question and resist them.

Principle 1: Enable Connections to Data

Our first principle suggests that algorithmic literacies can be supported by offering children opportunities to write programs that interact with data about their worlds. This principle stems from our experience with both Scratch Cloud Variables and Scratch Community Blocks. In both cases, we have found that even relatively simple connections to data from a programming toolkit enable scenarios where children ask questions about the algorithms that shape, store, and use information they create and care about.

We developed Scratch Cloud Variables as part of the second generation of the Scratch programming language (Scratch 2.0). The system allowed Scratch users to store values in variables in ways that persist beyond the run time of their program. Additionally, Scratch Cloud Variables is global in that everyone interacting with a project that uses Scratch Cloud Variables would see the same data (Dasgupta, 2013a). This support for persistent global data, combined with the fact that Scratch 2.0 projects were stored online and ran in a web browser, allowed for functionality in Scratch projects such as global high-score lists, surveys, collaborative art projects, and more (figure 6.1).

During beta testing of the system, children raised concerns about potential privacy threats and control over data made possible by the system. Because a Scratch project could store data persistently, it was possible to create relatively simple Scratch code that would ask for someone's Scratch username (e.g., for a Scratch "guestbook" project) and store it indefinitely. The only way to remove the data, once stored, was to ask the project creator to do so. A Scratch community member noted:

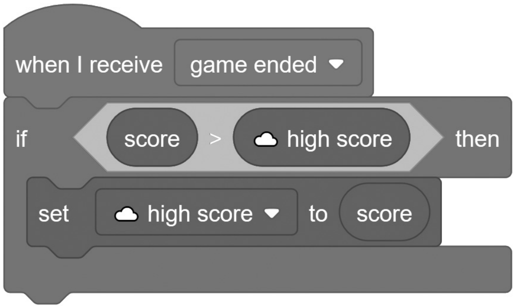


Figure 6.1
An example of a Scratch script using Scratch Cloud Variables. The script determines if the score of the concluded game is higher than the recorded high score. If so, the cloud variable high score (indicated by the cloud icon before the variable name) is updated.

> So if I typed in my first name [into the project] without thinking about it, after that everyone who views the project can see my name [. . .]. Furthermore to remove it I have to contact the owner of the project and request they remove it from the cloud data list. (Dasgupta, 2013b, p. 96)

This example shows that even relatively simple connections with data open up possibilities that enable Scratch community members to think about questions of algorithmic surveillance and power.

We developed the second system, Scratch Community Blocks, by adding programming blocks representing programming primitives into Scratch that could access public metadata about projects and users in the Scratch online community database (Dasgupta & Hill, 2017). An example of the system is shown in figure 6.2. For example, with Scratch Community Blocks, it was possible to create Scratch programs that would access how many times a Scratch project shared in the community had been viewed. Community-wide statistics, such as total number of registered users in the community, were also accessible programmatically through Scratch Community Blocks. These two sets of programming blocks were combined by a young Scratch user to make a project that calculated

Figure 6.2
Example code using Scratch Community Blocks. The code iterates through all the projects shared by the user scratchteam. In each iteration, the Scratch character being controlled by the code says the project title. Image from Dasgupta and Hill (2017).

what proportion of the broader Scratch community had viewed a given Scratch project.

Soon after this project was shared, Scratch community members began discussing the difference between views and viewers (i.e., a single user may view a project multiple times in ways that increase the project's view count). These results were confusing because, at the time, the Scratch website counted views using an algorithm that not only tried to count as many views as possible (e.g., from non-logged-in users) so that project creators would see that their creations had an audience, but also prevented community members from generating views synthetically (e.g., by repeatedly refreshing a project page).[1] Community members noted that one of the most popular projects on the site had a view count that exceeded the number of user accounts on the site.

> **Commenter A:** that's so cool! almost 0.5% of all the users on scratch have viewed my projects and that's a lot :B but crossstitch's[2] results are indeed slightly dubious . . . over 100% of people have viewed his projects which is awesome but impossible—love the project!! ^o^
>
> **Commenter B:** @CommenterA I think it's because its based on views, not each specific player.
>
> **Commenter A:** @CommenterB that's awesome :D people who haven't registered on scratch have viewed a significant amount of his projects yes

> **Project Creator:** @CommenterA Yeah what @CommenterB said is correct.
>
> (Hautea et al., 2017, p. 925)

Through these comments, users worked collectively to show how data are not objective but require *interpretation*, and that data generation is shaped by others' decisions.

Bowker (2005) has argued that "raw data is [. . .] an oxymoron" (p. 184). In an edited volume with the same name, Gitelman (2013) noted that "data are imagined and enunciated against the seamlessness of phenomena" (p. 3). Often, this imagination and enunciation materializes in an algorithm that collects data, such as the viewership statistics of Scratch projects. Enabling children to access that data through algorithms they implemented using Scratch Community Blocks led them to discover illuminating patterns in data, such as view counts of popular projects exceeding the number of community members. As in the extended example of the dialogue about how views are counted on the Scratch website, this in turn led children to attempt to reconstruct how data might have been imagined in the first place.

Resnick and Silverman (2005) posit that "a little bit of programming goes a long way," meaning children can combine relatively simple and limited programming constructs toward a broad range of creative outcomes. In our work, we see a similar phenomenon emerge where simple programming constructs, combined with data in straightforward ways, enable children to uncover structures and assumptions in algorithmic systems. This process allows them to question and discuss algorithmic data collection.

Principle 2: Create Sandboxes for Dangerous Ideas

Our second principle suggests that the development of critical algorithmic literacies can be supported through creating sandboxes for dangerous ideas. Like any sociotechnical tool, algorithms offer benefits and carry harms. We describe algorithms as "dangerous" to highlight the way they can have powerful, unanticipated, and negative consequences (Smith, 1985). For example, the algorithm behind a real estate search tool may allow the user to filter houses for sale by school rating, but it may not take the history of underfunding of schools in African American and low-income neighborhoods into account. In this way, the search algorithm might unintentionally become a way for potential house buyers to filter for affluent, predominantly White neighborhoods (Noble, 2018, p. 167).

With the deployment of Scratch Community Blocks, metadata about user accounts such as number of followers and number of projects were made programmatically accessible. These numbers can be used as proxies for measures of experience—that is, more projects or more followers indicates more experience with Scratch—but both are far from perfect measures. Although restricting interaction with a project to more experienced users might be an attractive feature to some Scratch users, using these measures as a gatekeeping mechanism can discriminate against newcomers. This was exactly the concern raised by a 13-year-old member of the Scratch community who noted that the algorithm to carry out such discrimination is trivial using Scratch Community Blocks and a single if statement:

> I love these new Scratch Blocks! However I did notice that they could be used to exclude new scratchers or scratchers with not a lot of followers by using a code: like this:
>
> ```
> when flag clicked
>     if then user's followers<300
>         stop all.
> ```
>
> (Hautea et al., 2017, p. 925)

Thus, this young user noted that algorithmic systems can be dangerous in that they can enable discrimination. This type of observation is far from uncommon among Scratch users, and it reflects a key step toward developing critical algorithmic literacies. That said, it is only possible because of the possibilities introduced by the system.

The notion of engaging with and exploring dangerous ideas is not new in education. Problematic theories are studied as a part of understanding history; discriminatory scenarios are analyzed as a part of engaging with the idea of justice; and potentially physically dangerous experiments are carried out in school and college chemistry laboratories. Although these activities all represent different types of danger, the pedagogical activities around them typically incorporate appropriate safety mechanisms. For the pedagogy of fields like chemistry, this is a topic of ongoing research and study (Alaimo et al., 2010).

An example from our own work that led us to consider the importance of making space for dangerous ideas is a feature introduced in Scratch 2.0—an username block that "reports" the username of the project's viewer if they are logged in (figure 6.3). The username block made new types of functionality possible, including a form of surveillance, by making it much easier to know who (in terms of Scratch usernames) had accessed a given project. As designers, we were also concerned that the block could be used for discrimination within a Scratch project (e.g., by disallowing certain usernames from playing a Scratch game) or to evade moderators (e.g., to have a Scratch project behave in a specific way only for known moderators in the community). On the other hand, we found that the block also made new conversations around surveillance, discrimination, data,

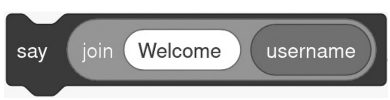


Figure 6.3
The username block introduced in Scratch 2.0. The block "reports" or returns the username of the person viewing the project, or it remains blank if the viewer is not logged in.

and algorithms possible. Achieving a balance between enabling exploration of dangerous ideas and potentially problematic applications of these ideas is not easy. This is especially the case among historically marginalized groups in STEM learning, who may be more vulnerable to discrimination and surveillance.

We only considered the feature because usernames in Scratch are, by community policy, not directly tied to identities in the real world. As a result, the consequences of username-based surveillance in Scratch would be less serious than surveillance of email accounts or other social media accounts. We also considered several approaches to making the username block not report the username directly. Many of these were technical. For example, an initial prototype of the block reported back an alphanumeric value that would remain consistent for a given user accessing a given project over time so that a user interacting with the project could not be identified by username[3] (Dasgupta, 2012). Partly because this idea was difficult to explain, we did not adopt this approach and reverted to the earlier, simpler approach of reporting the username. However, as an added measure, the Scratch project–viewing interface was modified so that it warned users about the block existing in a given project before they ran it and encouraged users to log out of Scratch if they wished to avoid being tracked by a project (Dasgupta, 2013b).

We also deployed the username block with considerable caution, carefully monitored its usage, and were ready to roll back the feature. The Scratch community has a complex and extensive moderation and governance infrastructure that has been described by Lombana Bermúdez (2017) as a combination of "proactive and reactive moderation . . . with the cultivation of socially beneficial norms and a sense of community" (p. 35). We felt that these structures had the potential to prevent and mitigate uses of the feature that could, in theory, go against the friendly nature of the community. In the design phase of the username block, we had many conversations with Scratch's moderation team and with children.

These conversations continued after the feature's launch so we could understand how the new block was being used by the broader community and adapt the system if needed.

To describe our approach of trying to balance exploring "dangerous ideas" and the potentially problematic applications of such ideas, we used the metaphor of sandboxes. Just as a sandbox allows children to play with earth in ways they couldn't for the rest of the playground, our design consists of creating clear boundaries and implementing sociotechnical strategies that prevent any use that might go against community values. The metaphor of a sandbox is common in the field of computer security, where untrusted applications are said to run in a "sandbox" isolated from unneeded resources and other programs (Schreuders et al., 2013). For example, a mobile phone sandboxing system might ensure that an app that does not need access to the camera does not have access to it. In the case of the username block, we spent several design iterations working with children and community moderators to ensure there were enough "walls" (e.g., warning messages in projects that use the block) before we felt we had achieved a balance between encouraging exploration and preventing potential violations of the Scratch community values. In computer and information security pedagogy, using sandboxes to allow learners to experiment with software vulnerabilities is an established practice (Du & Wang, 2008). Computer security researchers and educators have asked students to construct speculative fiction to engage with dangerous ideas and to imagine these ideas' impact on society (Kohno & Johnson, 2011). Our experience suggests that a similar approach may work for critical algorithmic literacies as well.

Principle 3: Adopt Community-Centered Approaches

Our third principle suggests that designers should incorporate community-centered approaches. These approaches would allow a design to leverage existing community values that an algorithm might change or challenge. This principle reflects increasing recognition of the importance of centering values in computing learning. For example, in a keynote presentation to the 2012 ACM SIGCSE Technical Symposium, Abelson (2012) called for a focus on "computational values," which he defined as commitments that "empower people in the digital world," and which he argued are "central to the flowering of computing as an intellectual endeavor." Justice, respect for privacy, and nondiscrimination are examples of such values.

Prior work in human-computer interaction has argued that values related to a computing system are "something to be discovered" in the context of a given community (Le Dantec et al., 2009, p. 1145). In turn, values can also influence the sense of identity of a learner within their communities. In recent work in the learning sciences, Vakil (2020) proposed the phrase "disciplinary values interpretation" to describe how learners seek to understand what a discipline being studied "is 'all about,' and what it might mean for them to be a part of it as they begin to imagine their future academic, career, and life goals" (p. 7). Vakil has also called for more understanding of, and attention to, "adolescents' political selves and identities, and how these identities become intertwined with learning processes" (p. 22).

In our work, we have seen the dynamic described by Vakil play out as children evaluate technological possibilities in terms of their values. For example, we saw that children using Scratch Community Blocks questioned algorithms by describing algorithmic outcomes that conflicted with the collective value of the Scratch community. Multiple community members expressed concerns about Scratch Community Blocks that enabled projects to rank community members based on the number of followers, and pointed out that this might shift the Scratch community's values from celebrating creativity and expression to emphasizing popularity. For example, a 12-year-old Scratch community member expressed concern that the new system allowed community members with a smaller number of followers to be made fun of through Scratch projects:

> [. . .] you can easily make fun of someone for example, "You only have 2 followers! Ha! Well I have 10!" (Hautea et al., 2017, p. 927)

Similarly, the aforementioned 13-year-old using Scratch Community Blocks who pointed out that code using the new system could be used to block newcomers from projects thought it was problematic because inclusivity is a core value of the Scratch community, and the algorithmic discrimination that this user correctly identified as being made possible by the system stood in contrast to this value.

One challenge with systems enabling possibilities that go against established community values is that many unsocialized newcomers do not share their new community's values. Zittrain (2006) noted this as a challenge with "generative" systems and platforms, where the outcomes made possible by the system were both positive and negative. With Scratch Cloud Variables, we recognized this issue and implemented a system where the Scratch Cloud Variables feature would only be available to users who were active in the community for some time (Dasgupta, 2013a; Dasgupta & Hill, 2018). By granting only socialized users access to the dangerous feature, we reasoned that newcomers would be allowed to learn Scratch's community values before they could receive access to features that enabled them to flout them.[4]

Our experience suggests that critical approaches to algorithms are driven by the values of the communities in which algorithms are enacted. Of course, communities vary in scope and character, and they can range from groups of friends to families, classrooms, and entire nations. In a 2014 report, the Executive Office of the President invoked values enshrined in the legal structure of the United States when it stated that "big data technologies" have the "potential to eclipse longstanding civil rights protections in how personal information is used in housing, credit, employment, health, education, and the marketplace" (Podesta, 2014, p. 3).

Of course, not all values are aligned with outcomes that educators seek to reinforce. Values frequently conflict with each other, and widely shared values can sometimes be problematic. Moreover, Benjamin (2019) describes how specific ways of imagining a value—a "master narrative" (p. 134)—can overwhelm alternatives. For example, Philip et al. (2013) draw from their classroom experience in a US public school to describe how the underlying assumption among students debating big data and its implications was "that students, particularly urban children of color, would academically, socially, occupationally, or politically benefit simply by virtue of exposure to big data and new technologies" (p. 117). Philip et al. explain that the design of their new curriculum did not consider the relative lack of opportunities for students of color, leading to an overwhelming focus on one particular framing of big data technology as an equalizer. This example serves as a warning for designers to carefully interrogate a range of values before designing.

Principle 4: Support Thick Authenticity

Finally, we argue that thick authenticity—a principle that applies to learning technology design in general—plays a crucial role in developing critical algorithmic literacies. Authenticity is a complicated concept in the context of learning. Although it is common to encounter terms related to the "real world" and "real-world problems" in popular and scholarly discourse on education, degrees and dimensions of "realness" vary enormously. For example, a learning exercise might involve a fictional scenario where a problem is real but a situation is not (Petraglia, 1998).

Ultimately, "realness" is determined by learners, and "real" learning experiences, problems, and metaphors may be unfamiliar to learners for individual or cultural reasons. Teaching probability using examples based on playing cards may lead to poor learning experiences if students have never played cards. As corollaries, stronger forms of authenticity emerge when learners have more say in the design and direction of their learning activities, and higher degrees of authenticity are associated with better learning outcomes. For example, while teaching Maori schoolchildren English, Ashton-Warner (1986) found that a compelling strategy to engage her students was to ask them to write about themselves, about their own stories, in their own words—a process she called "organic writing."

Questions of authenticity are likely to be relevant to the type of engagement necessary for supporting the development of algorithmic literacies in children. For example, a baseball analytics algorithm may generate critical engagement when the learner is a young baseball fan who is going to poke holes in the algorithm's assumptions. The same algorithm would likely elicit a lukewarm response, at best, from someone without an interest in baseball. To most learners outside the United States, learning activities that involve baseball are not suitable at all. In our design work, we have drawn inspiration from Ashton-Warner and asked what "organic writing" might look like for developing critical algorithmic literacies (Dasgupta, 2016).

We have also drawn from Shaffer and Resnick (1999), who describe "thick authenticity" as "activities that are personally meaningful, connected to important and interesting aspects of the world beyond the classroom, grounded in a systematic approach to thinking about problems and issues, and which provide for evaluation that is meaningfully related to the topics and methods being studied."

In our work with Scratch Community Blocks, children using the system engaged with complex ideas about power and algorithms because the data they were accessing, and the algorithms they were designing, reflected their experiences, friends, and community in Scratch. If the same interface within Scratch had provided access to nearly any other data source, it would have been less effective in promoting algorithmic literacy among the community of Scratch users to whom our system was deployed. Considering that most children do not use Scratch, the effectiveness of our systems is likely to be limited among most children.

That said, other contexts might present similarly promising opportunities. For example, families' interactions are increasingly shaped by algorithms and data inherent in smart home technologies. Although it is still less common among children, many people are increasingly collecting data about aspects of their personal lives through quantified-self approaches (Lee, 2013). Because algorithms are increasingly prevalent in a range of contexts, an increasingly wide range of settings offers rich opportunities for promoting algorithmic literacies through thick authenticity.

Discussion

In our own design experiences spanning many iterations, we encountered numerous tensions and open questions in terms of how to best engage the broadest possible set of Scratch community members in critiquing algorithms. The evidence emerging from our work suggests that certain design principles for computational construction kits may support the development of a range of critical algorithmic literacies. Our four design principles reflect a broader perspective—that young learners should go beyond simply observing algorithmic systems and be given opportunities to *create* algorithmic systems. We argue that when children take advantage of these opportunities, some will question algorithms in meaningful ways. In empirical work we conducted, we employed grounded theory (Charmaz, 2006) to analyze the discussions, comments, and activities of children engaging in creative design activities using Scratch Community Blocks (Hautea et al., 2017). Most of the examples we identified of children questioning algorithmic systems emerged from the process of active creation with toolkits.

Though our work is exclusively focused on the Scratch online community, we believe that the lessons we've distilled can be applied to other contexts. For example, the first author of this chapter (Dasgupta) used the "enable connections to data" principle in an introductory college-level Python course to illustrate how gender is often encoded as binary in software. As part of a relatively simple exercise that involved using conditional (if-else) statements, he asked students to use publicly available data on the recommended daily allowance of calcium to make an interactive tool that asked a few questions about an individual (e.g., age, gender) and then recommended a daily allowance of calcium for them.

Because the publicly available data table that is used for this purpose encoded gender as binary,[5] students ended up designing their programs with the inbuilt assumption of binary gender, with a small proportion of the students questioning the practice in their submission (e.g., as code comments). This provided Dasgupta with an opportunity to engage students—after they had written the program—in a discussion starting with the prompt, "What's wrong with this assignment?" Students pointed to the notion of a binary variable to represent gender, and this led to a broader discussion about the choices programmers make in modeling the world in their algorithms (Costanza-Chock, 2018; Smith, 1985).

Similarly, one might use the principle of "sandboxes for dangerous ideas" in a web programming exercise for college students by scaffolding an exercise around a survey system's design and conversations around the choices students make about tracking their survey-participants' identities. What is an acceptable technical solution to prevent repeat participation in an anonymous survey? Can a "technical solution" even exist? There are likely many other ways to help children understand and question algorithmic systems. For example, approaches such as co-designed games have yielded promising results for engaging children in understanding privacy (Kumar et al., 2018). Similar approaches may emerge toward other aspects of critiquing data and algorithm-driven systems as well.

Limitations

Our work is limited in that it has focused on individual learners and case studies. Our examples are also limited in that they show what is possible when our design implications are embraced, but not necessarily what is likely. For example, Scratch users—especially those who chose to engage with us by responding to our recruitment calls, coming to our workshops, and participating in other ways—are unlikely to represent children in general. We do not claim that the examples described here represent how most children would respond to the same tools and contexts. Indeed, we believe that the uniqueness of children's context, interests, and backgrounds likely means that no single tool or approach will work for all children.

Moreover, although several of our study participants questioned and critiqued algorithmic possibilities, many engaged with at least some of these problematic possibilities uncritically (e.g., by making Scratch Community Blocks projects that would only work for community members with more than five followers). More structured support, such as the "what's wrong with this assignment?" prompt in the classroom example we provided, will often be needed to engage a broader group in considering these questions.

In a sense, these facts remind designers and educators of the perils of a technocentric approach in learning—"the tendency to give . . . centrality to a technical object" (Papert, 1987, p. 23). Brennan (2015) offered a model of how a designer can support the work of educators toward using technology in the service of learning in a nontechnocentric way. Beyond systems and curriculum designed based on the four principles we have outlined, such models represent essential and complementary components that need to be in place in a learning environment for learners to develop critical algorithmic literacies. By stating these limitations and recognizing that this is a work in progress, we offer our principles in the hopes that other designers of educational technologies and experiences will build on our work and contribute to the larger project of fostering critical algorithmic literacies in children.

Conclusion

In their paper on designing construction kits for children, Resnick and Silverman (2005) present their final design principle as "iterate, iterate—then iterate again." They conclude by stating that this applies to their principles as well. Our four principles are no exceptions to this excellent advice. We intend to keep iterating on our principles, taking them apart, putting them back together, and changing them. We offer our principles with humility and a sincere desire to work toward the dual goals of supporting children in understanding the algorithmic systems that are increasingly shaping their worlds and doing what we can to give them the intellectual tools to question and resist them.

Acknowledgments

Many of the projects referred to in this chapter were financially supported by the US National Science Foundation (1002713, 1027848, DRL-1417663, and DRL-1417952). We would also like to acknowledge feedback and support from Mitchel Resnick, Natalie Rusk, Hal Abelson, Brian Silverman, Amos Blanton, and other members of the Scratch team. Finally, none of this work would have been possible without the children who generously tried out our

new technologies, gave us feedback, and inspired us in multiple ways with their ingenuity and kindness.

Notes

1. The first author of this article had implemented this particular algorithm for the Scratch website at that time.

2. crossstitch (username changed) was at that time the creator of some of the most popular projects on Scratch.

3. Internally, this alphanumeric value was being generated by one-way cryptographic hash of the username, project ID, and a cryptographic salt.

4. A second part of the reasoning is that if someone used Scratch Cloud Variables in a problematic way, they might get banned and lose access to the account into which they had poured substantial time and resources.

5. Current versions of the table and the associated text have discontinued the practice; earlier versions of the page can be accessed from the Internet Archive (e.g., https://web.archive.org/web/20200817125236/https://ods.od.nih.gov/factsheets/Calcium-HealthProfessional/).

References

Abelson, H. (2012). From computational thinking to computational values. *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 239–240. https://doi.org/10.1145/2157136.2157206

Agre, P. E. (2014). Toward a critical technical practice: Lessons learned in trying to reform AI. In G. Bowker, S. L. Star, L. Gasser, & W. Turner (Eds.), *Social science, technical systems, and cooperative work: Beyond the great divide* (pp. 131–157). Taylor & Francis. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1689049

Alaimo, P. J., Langenhan, J. M., Tanner, M. J., & Ferrenberg, S. M. (2010). Safety teams: An approach to engage students in laboratory safety. *Journal of Chemical Education*, *87*(8), 856–861. https://doi.org/10.1021/ed100207d

Ashton-Warner, S. (1986). *Teacher*. Simon & Schuster.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Bowker, G. C. (2005). *Memory practices in the sciences*. MIT Press.

Brennan, K. (2015). Beyond technocentrism: Supporting constructionism in the classroom. *Constructivist Foundations*, *10*(3), 289–296. http://constructivist.info/10/3/289.brennan

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

Costanza-Chock, S. (2018, July 16). Design justice, A.I., and escape from the matrix of domination. *Journal of Design and Science*. https://doi.org/10.21428/96c8d426

Dasgupta, S. (2012). *Learning with data: A toolkit to democratize the computational exploration of data* [Masters thesis, Massachusetts Institute of Technology]. DSpace@MIT. http://dspace.mit.edu/handle/1721.1/78203

Dasgupta, S. (2013a). From surveys to collaborative art: Enabling children to program with online data. *Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13)*, 28–35. https://doi.org/10.1145/2485760.2485784

Dasgupta, S. (2013b). Surveys, collaborative art and virtual currencies: Children programming with online data. *International Journal of Child-Computer Interaction*, *1*(3–4), 88–98. https://doi.org/10.1016/j.ijcci.2014.02.003

Dasgupta, S. (2016). *Children as data scientists: Explorations in creating, thinking, and learning with data* [Doctoral thesis, Massachusetts Institute of Technology]. DSpace@MIT. http://dspace.mit.edu/handle/1721.1/107580

Dasgupta, S., & Hill, B. M. (2017). Scratch community blocks: Supporting children as data scientists. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 3620–3631. https://doi.org/10.1145/3025453.3025847

Dasgupta, S., & Hill, B. M. (2018). How "wide walls" can increase engagement: Evidence from a natural experiment in Scratch. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, 361:1–361:11. https://doi.org/10.1145/3173574.3173935

D'Ignazio, C., & Bhargava, R. (2015, September). Approaches to building big data literacy. *Proceedings of the Bloomberg Data for Good Exchange Conference 2015*. https://dam-prod2.media.mit.edu/x/2016/10/20/Edu_D'Ignazio_52.pdf

diSessa, A. A. (2001). *Changing minds: Computers, learning, and literacy*. MIT Press.

Du, W., & Wang, R. (2008). SEED: A suite of instructional laboratories for computer security education. *Journal on Educational Resources in Computing*, *8*(1), 1–24. https://doi.org/10.1145/1348713.1348716

Freire, P. (1986). *Pedagogy of the oppressed*. Continuum.

Gitelman, L. (Ed.). (2013). *"Raw data" is an oxymoron*. The MIT Press.

Hautea, S., Dasgupta, S., & Hill, B. M. (2017). Youth perspectives on critical data literacies. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 919–930. https://doi.org/10.1145/3025453.3025823

Kafai, Y. (2006). Constructionism. In K. R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (1st ed., pp. 35–46). Cambridge University Press.

Kafai, Y., Proctor, C., & Lui, D. (2019). From theory bias to theory dialogue: Embracing cognitive, situated, and critical framings of computational thinking in K-12 CS education. *Proceedings of the 2019 ACM Conference on International Computing Education Research*, 101–109. https://doi.org/10.1145/3291279.3339400

Kohno, T., & Johnson, B. D. (2011). Science fiction prototyping and security education: Cultivating contextual and societal thinking in computer security education and beyond. *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education—SIGCSE '11*, 9. https://doi.org/10.1145/1953163.1953173

Kumar, P., Vitak, J., Chetty, M., Clegg, T. L., Yang, J., McNally, B., & Bonsignore, E. (2018). Co-designing online privacy-related games and stories with children. *Proceedings of the 17th ACM Conference on Interaction Design and Children—IDC '18*, 67–79. https://doi.org/10.1145/3202185.3202735

Le Dantec, C. A., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience: Evolving value sensitive design in support of value discovery. *Proceedings of the 27th International Conference on Human Factors in Computing Systems—CHI '09*, 1141. https://doi.org/10.1145/1518701.1518875

Lee, C. H., & Garcia, A. D. (2015). "I want them to feel the fear . . .": Critical computational literacy as the new multimodal composition. In *Gamification: Concepts, methodologies, tools, and applications* (pp. 2196–2211). IGI Global. https://doi.org/10.4018/978-1-4666-8200-9.ch111

Lee, C. H., & Soep, E. (2016). None but ourselves can free our minds: Critical computational literacy as a pedagogy of resistance. *Equity & Excellence in Education*, *49*(4), 480–492. https://doi.org/10.1080/10665684.2016.1227157

Lee, V. R. (2013). The quantified self (QS) movement and some emerging opportunities for the educational technology field. *Educational Technology*, *53*(6), 39–42. http://www.jstor.org/stable/44430216

Lombana Bermúdez, A. (2017). Moderation and sense of community in a youth-oriented online platform: Scratch's governance strategy for addressing harmful speech. In *Perspectives on Harmful Speech Online*. Berkman Klein Center for Internet & Society Research Publication. https://dash.harvard.edu/handle/1/33746096

Monroy-Hernández, A., & Resnick, M. (2008). Empowering kids to create and share programmable media. *Interactions*, *15*(2), 50–53. https://doi.org/10.1145/1340961.1340974

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Papert, S. (1987). Computer criticism vs. technocentric thinking. *Educational Researcher*, *16*(1), 22–30. https://doi.org/10.3102/0013189X016001022

Papert, S., & Harel, I. (1991). Situating constructionism. In *Constructionism* (Vol. 36, pp. 1–11). Ablex Publishing.

Petraglia, J. (1998). *Reality by design: The rhetoric and technology of authenticity in education*. Lawrence Erlbaum Associates. http://www.myilibrary.com?id=232329

Philip, T. M., Schuler-Brown, S., & Way, W. (2013). A framework for learning about big data with mobile technologies for democratic participation: Possibilities, limitations, and unanticipated obstacles. *Technology, Knowledge and Learning*, *18*(3), 103–120. https://doi.org/10.1007/s10758-013-9202-4

Podesta, J. (2014). *Big data: Seizing opportunities, preserving values*. Executive Office of the President.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60–67. https://doi.org/10.1145/1592761.1592779

Resnick, M., & Silverman, B. (2005). Some reflections on designing construction kits for kids. *Proceedings of the 2005 Conference on Interaction Design and Children*, 117–122. https://doi.org/10.1145/1109540.1109556

Schreuders, Z. C., McGill, T., & Payne, C. (2013). The state of the art of application restrictions and sandboxes: A survey of application-oriented access controls and their shortfalls. *Computers & Security*, *32*, 219–241. https://doi.org/10.1016/j.cose.2012.09.007

Shaffer, D. W., & Resnick, M. (1999). “Thick” authenticity: New media and authentic learning. *Journal of Interactive Learning Research*, *10*(2), 195–215.

Smith, B. C. (1985). The limits of correctness. *ACM SIGCAS Computers and Society*, *14*, *15*(1, 2, 3, 4), 18–26. https://doi.org/10.1145/379486.379512

Tygel, A., & Kirsch, R. (2016). Contributions of Paulo Freire for a critical data literacy: A popular education approach. *The Journal of Community Informatics*, *12*(3). https://ojs.uwaterloo.ca/index.php/JoCI/article/view/3279

Vakil, S. (2020). “I’ve always been scared that someday I’m going to sell out”: Exploring the relationship between political identity and learning in computer science education. *Cognition and Instruction*, *38*(2), 87–115. https://doi.org/10.1080/07370008.2020.1730374

Zittrain, J. L. (2006). The generative internet. *Harvard Law Review*, *119*(7), 1974–2040. http://www.jstor.org/stable/4093608

7

Authenticity and Co-Design

On Responsibly Creating Relational Robots for Children

Marion Boulicault and Milo Phillips-Brown
with Jacqueline M. Kory-Westlund, Stephanie Nguyen, and Cynthia Breazeal

Introduction

Meet Tega. Blue, fluffy, and AI enabled, Tega is a *relational robot*: a robot designed to form relationships with humans (figure 7.1 shows Tega interacting with a child). Created to aid in early childhood education, Tega talks with children, plays educational games with them, solves puzzles, and participates in creative activities like making up stories and drawing. Powered by AI algorithms, Tega adapts to each child's social, emotional, and curricular needs, thereby building a relationship that keeps them engaged and improves how they learn. For example, one of Tega's algorithms uses assessments of a child's language abilities to match the child with books to read that are at just the right language difficulty level to help build their vocabulary (Park et al., 2019).

For the past eight years, we at the Personal Robots Group at the MIT Media Lab have been developing and studying robots like Tega. Relational robots have the potential to play a part in addressing urgent social issues, such as ensuring access to high-quality early childhood education. We emphasize both *potential* and *play a part*. There is no guarantee that using relational robots will be either

Figure 7.1
A child with Tega.

effective or ethically sound, as is illustrated by analyses of recent proposals to use social robots for delivering therapy to children living with autism (McBride, 2020). And, like any technology, relational robots cannot address social issues on their own; they are sociotechnical tools that have the potential, under certain conditions, to contribute positively to broader interventions.

Our vision is for Tega to support parents, teachers, communities, and governments in helping children learn. So far, the results are promising. In working with over 400 preschool and kindergarten children (ages 4 to 6 years) in diverse public schools, we've found that children readily learn new words from robots like Tega (Kory-Westlund, Dickens, et al., 2017); emulate the robot's phrases and vocabulary during storytelling activities (Kory-Westlund, Jeong, et al. 2017; Kory-Westlund & Breazeal, 2019a); show more curiosity in response to a more curious robot (Gordon et al., 2015); and exhibit more creativity when the robot models creative behavior (Ali et al., 2019). Tega's relational nature has its own impact; the closer the relationships between child and robot, the more effectively the child learns (Kory-Westlund & Breazeal, 2019b). In one of our studies, for example, 49 children played one-on-one language-learning games with Tega once a week for eight weeks: the children who reported a closer relationship with Tega showed higher scores on language-learning metrics, such as vocabulary tests and ability to recall stories they'd heard or read (Kory-Westlund & Breazeal, 2019c; Kory-Westlund et al., 2018; Kory-Westlund, 2019).

Yet using relational robots for early childhood education raises pressing social and ethical issues. In designing relational robots for children, we are, in a sense, *designing relationships* between children and robots. If we are to design relational robots responsibly, then we must ask the questions: Should we be creating relationships between children and robots at all? And if so, what kinds of child–robot relationships should we design? That is, what should the relationship between a child and a robot like Tega be like? Or, what (if anything) makes a good child–robot relationship?

These questions have prompted discussion among our research participants and academics, and in the media (e.g., Turkle, 2007; Sparrow & Sparrow, 2006). Some of this discussion has centered on the concept of *authenticity*: good relationships are authentic relationships. A prominent concern is that child–robot relationships are inevitably inauthentic. That is, there is something inevitably inauthentic about *any* relationship that a child forms with a robot. If this is right, perhaps there is no way to responsibly design relational robots for children; perhaps we shouldn't be designing them at all (or, if we do, there must be *significant* benefit to outweigh the problem of inauthenticity). Here we explore this concern of authenticity.

It's important to emphasize that authenticity is far from the only salient ethical issue when it comes to relational robots. Others include concerns about data collection and ownership; privacy and security; social injustices concerning access to technology, corporate power, and the future of work; and aforementioned concerns about technological solutionism, to name just a few. Determining whether and how to design relational robots for use in children's education will require grappling with all these issues (and their intersections) in tandem.

The structure of this chapter is as follows. We begin in the next section by explaining what we mean by a "relational robot" and expand on the motivations for building relational robots to use in early childhood education. In the third section, we analyze two different concerns about authenticity. Our analysis draws on our group's empirical research as well as on insights from philosophy and disability studies. In the fourth section, we suggest a way forward. We argue that in order to design relational robots responsibly, it is ethically imperative that designers employ what is known as *co-design*, a framework that enlists stakeholders such as parents, teachers, and children themselves in answering the question: "What kinds of child–robot relationships should we design?" Using examples from our own research, we illustrate the significance of co-design for creating relational robots for children.

What Are Relational Robots? And Why Build Them?

What Are Relational Robots?

Relational robots, as we've said, are robots designed to form relationships with humans. They belong to a broader class of *relational technologies*, technologies that are designed to build relationships with humans. This use of the term *relational technology* dates back at least as far as Bickmore and Picard (2005).

The idea that humans have relationships with technologies like robots is based on an understanding of a relationship—endorsed by various scholars[1]—on which humans can form relationships with both humans and nonhumans (with pets, for example). This understanding of a relationship can be made clearer by considering a related concept, that of a *social interaction*. A social interaction is commonly understood as an interaction between two agents whose behaviors are interdependent; the actions of one agent are responsive to the actions of the other (Berscheid & Reis, 1998). Social interactions include behaviors such as conversing, meeting another's gaze, taking turns, displaying emotion, gesturing, and performing what's known as behavior mirroring—matching one's behavior to that of the other. (The behaviors that make up social interactions are known as *social behaviors*.) Many modern technologies engage in social interactions with humans—for example, entertainment robots like Aibo; personal home robots like Buddy, Jibo, and Mabu; and digital assistants like Alexa and Siri.[2]

Tega, too, socially interacts with humans—indeed, Tega is programmed to engage in a wide range of social behaviors. For example, Tega converses (using automatic speech recognition and by playing back recorded speech); meets the gaze of humans (e.g., Tega will "look" at the child's face when the child looks at it); and mirrors behavior (e.g., Tega will match the cadence of a child's speech or mirror a child's facial expressions). In our research, we've found that children tend to respond in kind. They readily converse with robots like Tega; mirror their behavior (e.g., mimic a robot's facial expressions); take turns; share information about themselves; and help robots during joint activities (e.g., they turn the pages of a digital storybook for the robot and help the robot "practice" storytelling by retelling stories) (Kanda et al., 2007; Kory-Westlund, 2019; Kory-Westlund & Breazeal, 2019a; Kory-Westlund et al., 2018; Park & Howard, 2015; Park et al., 2019; Serholt & Barendregt, 2016; Singh, 2018) (see figure 7.2 for an image of a child with Tega).

It takes more than just having a social interaction to be in a relationship, according to the understanding we are adopting. For example, if you meet the gaze of someone you pass on the street, you do not thereby have a relationship with them; or if you ask Alexa what the weather will be tomorrow and you get a response, you do not thereby have a relationship with Alexa. Rather, relationships require a series of repeated, personalized social interactions

Figure 7.2
School child posing with Tega.

that can elicit feelings of *responsiveness* and *commitment*, as they're known in the literature.

Relationships unfold over time: in a relationship, repeated social interactions inform future social interactions. Think of how your social interactions with a longtime friend differ from those with a stranger; this difference is partly due to a store of shared experiences. In a relationship, you can refer back to activities shared in the past. Or, when you respond to someone, or some*thing*, with whom you have a relationship, you can in a sense *personalize* your response based on what you know from past interactions. As we noted in the introduction, this is precisely what Tega does. That is, Tega uses AI technology to tailor its future social interactions with a child based on past interactions—for example, by picking books to read with children based on what it has learned about the child's literacy skills.

Feelings of responsiveness and commitment are umbrella terms that include positive feelings such as rapport, closeness, and attachment. Robots, of course, do not have feelings of responsiveness and commitment toward children; they do not have feelings at all. But children do. We've found in our studies that, for example, children report feeling as close to the robots as they feel to pets and favorite toys (Kory-Westlund et al., 2018). They readily say the robots are their friends (e.g., Kory, 2014) and frequently smile, laugh, and display various positive facial expressions when learning and playing with the robots (e.g., Kory-Westlund & Breazeal, 2019a; Kory-Westlund, 2019).

Perhaps you are still skeptical that the word *relationship* aptly describes the connections between children and relational robots. We explore skepticism of this kind in the subsection "Inauthenticity as Unreality?" Ultimately, though, it is not essential to our purposes that child–robot relationships deserve the name. What is important is that children interact with robots in certain ways, and conceive of them in certain ways, that are similar in some respects to how they interact with and conceive of humans. It is the ethical dimensions of these connections—not the label *relationship*—that is our concern.

Why Build Relational Robots for Early Childhood Education?

Improving the quality and equity of early childhood education for all children is an issue of US national educational importance (Hart & Risley, 1995; Garcia & Weiss, 2017). Early childhood is a critical time. It is when learning is most malleable and investments are most cost effective for spurring long-term benefits to cognitive, academic, behavioral, and socioemotional outcomes (Heckman et al., 2010). A child who cannot read adequately in the first grade has a 90 percent probability of reading poorly in the fourth grade and a 75 percent probability of reading poorly in high school (Torgesen, 2004). Tragically, about one-third of US children do not have access to high enough quality early childhood education programs to prepare them to meet standards for kindergarten entry (Torgesen, 2004).

We at the MIT Personal Robots Group are experimenting with technologies like Tega in the hopes of helping to address some of these pressing social and educational issues facing our youngest learners. As we mentioned in the introduction, Tega is designed to help young children develop language and literacy skills and improve key learning attitudes, such as curiosity, creativity, and the development of a growth mindset (the idea that one can develop talents and abilities through perseverance and effort [Dweck, 2008; Park et al., 2017]). Some evidence suggests that using AI technology to facilitate relationship-building between Tega and individual children makes Tega more effective at meeting these goals when compared to nonrelational technologies.[3] As such, Tega may be well positioned to support teachers in the classroom. For example, Boston-area preschool and kindergarten teachers from both private and public schools tell us that they would be excited to use robots like Tega during what they call "choice time"—a special time each day when children pick from a menu of different learning activities (Kory-Westlund et al., 2016).

Tega may also be effective at supporting parents and guardians with at-home learning (which became particularly urgent during the COVID-19 pandemic). For example, research shows that children benefit from responding to dialogic questions—that is, open-ended questions without clear right or wrong answers (Hargrave & Sénéchal, 2000; Valdez-Menchaca & Whitehurst, 1992; Whitehurst et al., 1988). Tega is programmed to ask dialogic, reciprocal questions as a parent reads a book to a child, supporting the parent in teaching their child (Boteanu et al., 2016; Chang et al., 2012; King, 1990; Nuñez, 2015).

Of course, issues concerning underfunding, support for teachers, and equitable access to high-quality early childhood education are complex social issues that will never have a purely technical solution. Indeed, a simplistic vision of "technological solutionism" diverts attention from the social dimensions of a problem and, more fundamentally, fails to recognize how the social and technical are inevitably intertwined (Šabanović, 2010; Winner, 1980; Morozov, 2013; Ames, 2019; McBride, 2020). However, when conceptualized, designed, and implemented responsibly as parts of broader sociotechnical interventions (as we discuss later), social robots have the potential to serve as tools for addressing urgent social problems.

Concerns about Authenticity

To design relational robots for children is to design relationships between robots and children. Therefore, responsibly designing relational robots requires us to address the ethically weighty question: "What kinds of relationships should we design?"

As noted in the introduction, one widely held answer to this question is based on authenticity: good relationships are (among other things) *authentic* relationships, so it is important that we design technologies for children that help create authentic relationships. During our studies, parents and teachers frequently raised the concern, in some form, that it's not possible for children to form authentic relationships with robots. This concern is echoed in the academic literature on relational robots: sociologist Sherry Turkle, for example, insists that, in contrast to authentic human-human relationships, human-robot relationships are "superficial," "pretend," and "inauthentic" (Turkle, 2007). Philosophers Robert Sparrow and Linda Sparrow (2006) contrast human–robot relationships with "genuine" human–human relationships.[4]

In this section, we analyze these concerns about authenticity. Our analysis reveals that there is no *one* unique authenticity concern; different ethical concerns go under the banner of "authenticity." We focus on two such concerns: the first is that child–robot relationships are *not real relationships*, and the second is that these relationships are *deceptive*.

A note on the scope of our ambitions. First, we aren't aiming to analyze all possible concerns about authenticity. Others don't relate to either reality or deception. For example, Turkle (2007) argues that another reason that human–robot relationships are ethically alarming is that they may, in time, lead children to form inauthentic *human–human* relationships. Second, we are not advancing an analysis of what authenticity *is per se*. Rather, we aim to analyze two often-voiced concerns about child–robot relationships—concerns that have been stated in terms of authenticity—to better understand how to responsibly design relational robots. And finally, as already emphasized, authenticity is but one of many complex, interconnected social and ethical issues that must be addressed in designing and using social robots in early childhood education.

On Theorizing about Authentic Connections

Before investigating concerns about the authenticity of child–robot relationships, we'd like to step back and comment on theorizing about the authenticity of connections between humans and non-humans more broadly.[5] It is strikingly easy to make unjustified and potentially harmful assumptions about the inauthenticity of such connections—a fact that comes into relief with an example from disability studies.

Theologian Julia Watts Belser (2016) highlights a common assumption about the connections between persons with disabilities and assistive technologies, like wheelchairs: they are thought of as a burdensome reliance, detracting from quality of life. Watts Belser illustrates this by pointing to the widely used phrase "wheelchair-bound," which evokes the idea of a wheelchair as something that "binds, traps, and constrains the human within its medicalized embrace" (2016, p. 6). In this view, people with disabilities would be better off if they didn't have to rely on assistive devices.

Watts Belser's own experience as a wheelchair user challenges this conventional thought. Rather than a burdensome reliance, she sees her connection with her wheelchair as one of "intimate engagement between wheel and flesh that is central to my own embodied experience" (p. 7). The blogger Wheelchair Dancer echoes Watts Belser in describing her own connection with assistive devices: "My crutches are part of my arms—when I use them to make a dance line—and extra spines when I use them to support me and when I shift all of my weight on to the conjunction of arm and crutch." Wheelchair Dancer argues that we should conceptualize "disabled anatomy not as a set of functioning and failed body parts, bits that have partially been replaced by technology, but as a body that is extended and expanded by its technology" (Watts Belser, p. 12). The connection between Wheelchair Dancer and her assistive technology is extensive, expansive, and empowering.

Once we consider Watts Belser's and Wheelchair Dancer's perspectives, it's hard to think of an adequate definition of authenticity that would label their connections with their wheelchairs and crutches as inauthentic. And yet this is the opposite of what we would expect if we adopted the conventional—and, to many, seemingly obvious—understanding of how persons with disabilities relate to assistive technologies, an understanding that is based on problematic ableist assumptions.

Of course, the relationships between children and robots are both practically and ethically different in significant ways from the connections between persons living with disabilities and assistive devices. Children don't, for example, usually think of robots as extensions of their bodies. And while child–robot relationships may face a certain stigma, that stigma cannot be compared to the ableist oppression that persons with disabilities face. Nonetheless, a lesson can be drawn from scholars working in disability studies: if we're theorizing about what counts as an authentic connection or relationship, we must be epistemically humble, which is to say that we cannot put too much weight behind our own thoughts and intuitions. We must look to those who have direct knowledge—or what's known as "lived experience"—of the connection or relationship. The judgments that may come easily must be carefully critiqued and interrogated. We ought to take extra caution with new types of relationships, like relationships between children and AI-enabled relational robots, where conventional wisdom may not apply.

Inauthenticity as Unreality?

With that in mind, let us turn to the concerns raised about the authenticity of child–robot relationships. In our research, we've found that when some study participants—such as teachers and parents—express concerns about authenticity, they sometimes seem to be expressing a concern that the relationship a child forms with a robot is somehow *unreal*, or at least *less real*, than the relationship a child forms with a teacher or friend. One could reconstruct this concern as follows. Human–human relationships are real; indeed, human–human relationships set the ideal for what a real relationship is. Any relationship that lacks the qualities of human–human relationships is a mere approximation of a real relationship. It is less than real, and therefore inauthentic.

This thought has intuitive appeal. Although human–robot relationships have various qualities found in paradigmatic human–human relationships (see previous section, "What Are Relational Robots? And Why Build Them?"), they lack many others. Today's robots do not empathize with a child who has stubbed her toe; they do not feel joy if a child writes them a thoughtful note; they do not care if they never again see a child with whom they've interacted. One would be quick to label "inauthentic" human–human relationships that lack these qualities: imagine someone who claims to be your friend but who doesn't empathize with you, is not moved by a thoughtful note, or wouldn't care if they never saw you again; this is *not a real friend*.

But we suggest that it is hasty to leap to the conclusion that *any* kind of relationship—especially human–nonhuman relationships—is fake or unreal if it lacks certain qualities, such as the ability to empathize. Is your relationship with your dog, for example, not real if he is indifferent to a thoughtful note? Presumably not. Human–human relationships don't set the standard for *all* relationships. Rather, we propose, there are relationships of different kinds, each of which might have different standards of "realness" or authenticity. What makes your relationship with a friend authentic, for example, is not, intuitively, the same as what makes your relationship with your dog authentic.

If this idea is right, then human–robot relationships may constitute "real" relationships—just a different kind of real relationship than human–human relationships. We've observed evidence of this in our research. We found that children generally do not conceive of robots as equivalent to their human peers and caregivers, or even as the same as their pets, toys, or computers (Kory-Westlund, 2019; Kory-Westlund & Breazeal, 2019a; Kory-Westlund et al., 2018).

This finding is well illustrated by a study we conducted to gauge how children perceive Tega. We asked children to complete a sorting activity in which we presented pictures of different entities, including a frog, a cat, a baby, a robot from a movie (like R2-D2 from the *Star Wars* films), a mechanical robot arm, Tega, and a computer (Kory-Westlund & Breazeal, 2019c). Children were asked to place these pictures on a spectrum with a human adult on one extreme and a table on the other. Children frequently placed Tega near the middle, between a computer and a cat, indicating that they saw Tega as more humanlike than a computer but less humanlike than a cat (which they generally placed closer to the adult than to Tega). In another study—which we referenced in the subsection "What Are Relational Robots?"—we asked children to talk about how close they felt to Tega in comparison to pets, toys, friends, and parents. On average, children said they felt similarly close to Tega as to their pets and favorite toys, but less close than how they feel to friends and parents (Kory-Westlund et al., 2018). These data lend credence to the thought that child–robot relationships needn't be, or needn't necessarily be, a less real, approximate version of human–human relationships. Child–robot relationships may simply be a different kind of relationship, with their own distinct standards of "realness."

In other words, we're suggesting that just because child–robot relationships lack qualities of human–human relationships does not mean—as some have worried—that child–robot relationships are less real and therefore inauthentic. There is evidence, for example, that children consider robots a different kind of entity than humans, suggesting that child–robot relationships may likewise be of a different kind than human–human relationships. Child–robot relationships may have their own distinct standards of realness and authenticity. As such, it does little to simply charge that child–robot relationships are "unreal" without specifying a standard of "realness" or "authenticity" against which to judge the relationships.

Nonetheless, we don't think that the inauthenticity-as-unreality concern is misguided. The issue is how it has been *expressed*. When theorists and our research participants say they are concerned about unreality, we think they are most charitably understood as giving voice to a different concern: that child–robot relationships are somehow *off* or *not quite right*. In other words, child–robot relationships are—for a reason not so easily articulated by unreality—not the kinds of relationships we should be designing for our children. (It is not only unreal relationships that are problematic. Think, for example, of a child's relationship with a bully: this isn't a relationship that a child should be in, but that has nothing to do with unreality. It may be all too real!)

The inauthenticity-as-unreality concern seems to bring us right back where we started: "What kinds of child–robot relationships should we design?" (Or should we even be designing them at all?)

Inauthenticity-as-unreality doesn't help answer this driving question, since it doesn't say what standards of "realness" we should be judging the relationships against. In the section "Responsible Design with Authenticity in Mind: An Argument for Co-Design," we will address this driving question in a way that we argue is more effective than considerations of realness. But before that, we first consider another commonly raised concern about the authenticity of child–robot relationships, this one having to do with deception.

Inauthenticity as Deception?

According to a second authenticity concern, child–robot relationships are inauthentic not because they are unreal, but because they are *deceptive*. Some relational robots are programmed to represent themselves—in some sense or another—as empathetic, curious, or having several emotions or mental states. For example, Tega can mirror children's facial expressions, giving the appearance of an emotional reaction; or, when playing a learning game, Tega can say things like "Ooh!" while leaning forward and opening its eyes wide, giving the appearance of curiosity. Other robots we've designed, such as Green the DragonBot, explicitly ascribe themselves emotions, saying, for example, "I like playing with you!"

The concern is that in behaving in these ways, robots—or, more accurately, the robot's designers and programmers—may lead children to wrongly believe that the robots are capable of emotion (Picard & Klein, 2002; Sparrow & Sparrow, 2006; Turkle, 2007). This *inauthenticity-as-deception* concern can be understood in various ways (see Coeckelbergh, 2012, for a taxonomy of these various ways). Here, we articulate one version of the concern.

The idea that deceptive relationships are inauthentic is familiar from everyday life. If you learned that your partner has lied to you for decades about their real name; pretended to love you when they did not; or only cared about your relationship insofar as it served their professional aims, all of this not only would be hurtful but would indicate something about the relationship itself, too. A relationship built on deception can rightly be called inauthentic, at least to some extent and in certain cases.

Are children wrong about what robots are like? The concern that child–robot relationships are deceptive presupposes that children are indeed mistaken about what robots are like. But are they? Do children mistakenly believe that today's relational robots—like Tega—have attributes, like a capacity for emotion, that they do not in fact have?

Children *do* ascribe emotions to relational robots. They say things about robots like "She's kind," "if you just left him here and nobody came to play with him, he might be sad," and "he likes sharing stuff, like stories" (Kory-Westlund et al., 2018). One child, when asked what he would do if one of our robots was sad, suggested he would "buy ice cream to make him happy, robot ice cream" (Kory, 2014). But of course these robots lack the capacity to feel kind or sad; they lack the capacity to like; if they were given ice cream—whether robot or human ice cream—it would not make them feel anything at all.

One conclusion to draw is that children are indeed mistaken about what robots are like. We would like to counsel caution about accepting this conclusion too readily. First, as we noted in the subsection "Inauthenticity as Unreality?," children tell us that they don't conceive of robots as equivalent to friends, parents, or other humans. This may suggest that while children use words like "sad" to describe robots, they may conceive of the sadness that they ascribe to robots differently than the sadness they would ascribe to a friend or parent. Just as a child conceives of a robot eating "robot ice cream" rather than "human ice cream," so too might the child think of a robot as having "robot feelings" rather than "human feelings."

Second and most obviously, it's uncontroversial that children engage in make-believe games and play activities where they knowingly pretend that things are other than what they are. This is something adults do with children—pretending, for example, that a Winnie the Pooh bear or Furby is alive and has feelings. All of this is considered an important and positive childhood activity. It's not a stretch to see Tega playing a similar role to these toys. Indeed, we've found in our research that parents and teachers pretend that Tega has feelings. Given that children aren't "deceived" by a Winnie the Pooh bear or Furby, we shouldn't be too quick in concluding they're deceived by Tega.

What do inauthenticity-as-deception concerns mean for the design of relational robots? One could nonetheless argue that Tega and toys like Winnie the Pooh and Furby differ when it comes to deception. Tega does many things that such toys do not, like sustain conversations with children and match their facial expressions and the pace and cadence of their speech. And most distinctively, Tega collects data from children and uses AI technology to personalize and adapt its interactions over time. As this AI technology advances, it is easy to imagine that Tega-like robots of the future will behave in ways that leave children genuinely believing that robots have thoughts and feelings.

If this is the case now or in the future—that is, if child–robot relationships are or will be somehow deceptive—would that be a cause for concern? We'll argue that the answer to this question is not straightforward.

Adults frequently deceive children—or don't disabuse them when they're mistaken about certain things, like whether their pet has died, whether the tooth fairy exists, or whether their dinner contains vegetables. The ethical implications of such deception differ considerably from deception toward adults. Compare a parent sneaking vegetables into a child's dinner and telling them there are no vegetables versus a company doing the same with their employees. We may imagine that in both cases, the deception leads to an outcome that benefits the deceived; with the child and parent, though, the deception has a different moral complexion than with the employee and company.

Using relational robots does not, as we see it, raise some *distinct* or *new* concern over and above those about deception of vegetables not being in dinner or the existence of imaginary beings. Rather, it seems clear that in general, parents, teachers, and other caretakers can use limited deception for the benefit of children—that is, deception in select cases and to select ends. And using relational robots promises to be of the exact kind that warrants such limited deception: helping the child to develop intellectually and emotionally. As we noted previously, our research indicates that relational robots indeed help children learn.

More generally, deception seems to fall into a broad category of behavior whose moral status depends on whether the recipient is an adult or a child. While in many cases it would be wrong to *control* the lives of adults—for example, deciding what they eat, who they can socialize with, or what their bedtime is—such treatment is not only appropriate for children but also the responsibility of caretakers. Deception is a certain way of controlling a person.

This is not to say that *all* control of children is good; and in particular, not to say that all deception of children is good. Far from it. Our point is rather that the moral import of deceiving children is complex. With children, we cannot simply equate "a deceptive relationship" with "a relationship a child should not have" (nor can we equate it with "a relationship a child *should* have"). To evaluate the ethical import of deceiving a child, we need to know more, as philosophers have argued. In particular, we need to know the *context* in which the deception is taking place. For instance, we need to know *why* the child is being deceived (see, e.g., Pallikkathayil, 2019). Is it to facilitate learning? To eat more vegetables? To spend more money on toys? And *who* is doing the deceiving? (See, e.g., White, 2021.) A parent? Robot? Teacher? Corporation?

Recall that the overarching question that needs an answer is, "What kinds of relationships should we design?" According to the most straightforward understanding of the inauthenticity-as-deception concern, any deceptive relationship is problematic; if child–robot relationships are deceptive, that is automatically cause for concern. But as this subsection shows, things are not so clear-cut. Some deceptive relationships may be problematic, while others may not be. Simply pointing to deception (just like simply pointing to the notion of unreality) is insufficient for telling us which relationships we should design. To tell whether deception in a child–robot relationship is problematic, we need to know the context—the *who*, *when*, and *why* of the deception. This is all to say that we need to know the context surrounding the child–robot relationship to determine what kinds of relationships we should design.

Responsible Design with Authenticity in Mind: An Argument for Co-Design

We've said that in designing relational robots for children we are, in effect, designing relationships. This is because children will form different kinds of relationships with different kinds of robots. For example, whether a robot says that it feels certain ways, or how it responds to a child asking, "Do you love me?" may affect whether the relationship is deceptive (and thus, according to some, inauthentic).

In the previous section, we argued that the two authenticity concerns we considered don't take us far enough in determining the kinds of child–robot relationships we should design, or whether we should be designing such relationships at all. In this section, we offer a more promising path forward. Rather than aiming to identify a fixed definition of the kind of child–robot relationship we should be designing (e.g., giving a definition of an authentic relationship), we focus on the *process* by which we answer the question, "What kinds of child–robot relationships should we design?" More specifically, we'll argue that this question can be answered responsibly only if it is answered collaboratively, using a family of methodologies known as collaborative design, or *co-design*.[6]

"What Is Co-Design?" explains the spirit and methods of co-design. "The Case for Co-Design in Building Relational Robots for Children" argues that co-design is imperative for addressing the question, "What kinds of child–robot relationships should we design?" And "An Example of Co-Designing Relational Robots with Diverse Stakeholders" shows co-design of child–robot relationships in action: we describe how we at the MIT Personal Robots Group have used co-design methods in designing our relational robots.

What Is Co-Design?

Co-design, most simply, is design in *partnership* with the people and communities who are or might be affected by a given technology. As is common, we'll call these people and communities *stakeholders*. Co-design overlaps with related approaches known as *participatory design*, *human-centered design*, and *inclusive design*; and indeed, it is often used as an umbrella term for these approaches. Costanza-Chock (2020) offers a useful encapsulation of co-design as "the full inclusion of, and accountability to, and control by, people with direct lived experience of the conditions [that] designers . . . are trying to change" (p. 26). And Also Too, a design studio dedicated to co-design, describes their work as "guided by two core beliefs: first, that those who are directly affected by the issues a project aims to address must be at the center of the design process, and second, that absolutely anyone can participate meaningfully in design" (And Also Too, n.d.).

What does it mean to design in partnership with stakeholders? To answer this question, it is helpful to contrast co-design with *user research* methods, which aim to obtain information from stakeholders. For example, a designer creating a meditation phone application might conduct focus groups with potential users to learn what these stakeholders want and how they might interact with such an application. User research methods provide information, but it is up to the designers to determine what they will do with that information. For example, the application designers might use what they learn to ensure that the app helps users meet their own meditation goals. Or they might use the information to design the application to maximize the time a user spends on it, regardless of the user's goals and values.

Co-design is different. While user research methods might form an important *part* of a co-design approach, these methods alone are not sufficient for co-design. This is because co-design requires that stakeholders be included not only as *sources* of information but also as *decision makers*. If we were using co-design to design a meditation app, stakeholders would not only provide information to the designers; they would also be partners in making design decisions.

There is no one-size-fits-all approach to co-design; rather, co-designers use a variety of methods and strategies for including stakeholders as design partners, depending on the nature of the project and on the specific stakeholders. These might include participatory technology assessments (Banta, 2009; Hennen, 2012), citizen juries (Gooberman-Hill et al., 2008; Street et al., 2014), and global interdisciplinary observatories (Hurlbut et al., 2018). (For more details on these methods, see Sample et al., 2019.) There are also co-design methods specifically targeted toward children. Druin (2002), for example, articulates a framework where children can take a variety of roles in the broader design process of new technologies—that of user, tester, informant, or design partner. This framework emphasizes that all partners "must acknowledge that a child has the right to partake and possess an active role" in the design process.

Co-design is not new to the design of relational robots. Researchers like Selma Šabanović have argued for similar participatory approaches (Šabanović, 2010). A research team at the University of California San Diego used co-design methods in designing robots for dementia caregiving. They conducted a six-month community design–research process, built relationships with members of local community centers, and empowered caregivers by collaborating with them in designing physical prototypes (Moharana et al., 2019). Other research teams have adopted co-design methods in designing relational robots for children. For example, researchers have explored using cooperative inquiry methods with intergenerational teams in designing social robots for children (Arnold et al., 2016). This approach allows groups of children across age ranges, with different levels of knowledge and learning styles, to explore new information together. Researchers in the Netherlands and the United Kingdom working on designing robots for children with autism implemented co-creation sessions with children, family members, and professionals affected by autism spectrum disorder (Huijnen et al., 2017). To facilitate collaboration and trust among participants, the sessions were held in environments familiar to participants, who sat in a "U-shape" arrangement (as opposed to rows, for example) so they could look at each other while speaking.

The need for facilitating trust brings up one of the central challenges—and promises—of co-design. We live in a world with extreme social inequities and hierarchical power structures, illustrated forcefully by the growing power divides between the technology sector and the rest of society. It may be difficult to find ways to effectively include stakeholders as partners, especially those who have been historically excluded from design processes, such as those from low-resourced or otherwise marginalized communities. For instance, in the context of relational robots for children, family members from low-resourced communities may not have access to transportation or have the time or resources to attend co-creation sessions or lab meetings. In addition, stakeholders from marginalized groups may not trust the universities or corporations building these technologies. This is why a co-design approach requires accounting for stakeholder histories and power dynamics.

The Case for Co-Design in Building Relational Robots for Children

Why is co-design necessary for designing relational robots responsibly? As we just discussed, co-design says that to responsibly design any given technology, the design process must include those people and communities who are affected by the technology. The primary motivation behind co-design is a matter of *justice*: those affected by a technology deserve a say in how they will be affected. In other words, stakeholders of any given technology deserve a say in how that technology is designed (see, e.g., Costanza-Chock, 2020). We'll argue for something more specific: that stakeholders of relational robots deserve a say in answering the question, "What kinds of child–robot relationships should we design?"

Outside the context of relational robots, the question of what kinds of relationships children should have is the province of parents, teachers, children themselves, caregivers, communities, and so on—or rather it is their province within certain bounds. It is not the province, or not the sole province, of traditional designers of technologies. Why would things be any different with the question of which relationships children should have with relational robots? As co-design dictates, a broad range of stakeholders—not just product designers and researchers—need power over decisions about the kinds of child–robot relationships that children have.

To make the point more concrete, think about one of the authenticity concerns we examined earlier in "Concerns about Authenticity"—specifically, that child–robot relationships are deceptive (and thereby inauthentic). We argued that simply knowing that a child–robot relationship is deceptive (if it's deceptive at all) isn't enough to determine whether it's a relationship that children should or should not have. Deception may be problematic in certain contexts but not in others. One determining factor in whether deception is problematic is *who decides* to deceive the child. Imagine that a food corporation creates a snack for children without disclosing to the public that it contains vegetables. Imagine further that your child buys this snack and eats it. She has been deceived, and, it seems, in a problematic way. The problem is not that it's never okay to deceive children about the contents of their food. It could be fine for *you* (the parent) to trick your child into eating vegetables. Rather, the problem—or at least part of the problem—is that it is not the place of a corporation to decide on its own whether to deceive children. As a parent, you deserve to have a say in whether your child is deceived. A similar thing goes, we maintain, for if and when child–robot relationships should be deceptive. It is not the place of traditional designers to decide this matter alone; parents and other stakeholders deserve a say, too.

We don't mean to suggest that if parents, teachers, or other stakeholders think it's appropriate to deceive a child, then they are thereby correct. There are, as we've said, simply cases where children should not be deceived (for example, if parents deceive their children without regard to their interests). More generally, there are certain kinds of relationships—for example, abusive or oppressive relationships—that children should *never* have, regardless of whether parents, teachers, a community, or anyone else thinks they should. This sets a certain boundary on what child–robot relationships we should be designing. But within this boundary, the question remains: "What kinds of child–robot relationships should we design?" This question, we've argued, is for co-design to answer.

An Example of Co-Designing Relational Robots with Diverse Stakeholders

We have outlined the concept of co-design, and we have argued that co-design methods are imperative for the responsible design of relational robots for children. In this subsection, we offer an example, based on our work designing Tega and Green the DragonBot, of what it looks like in practice to apply co-design methods to the design of relational robots for children (see figure 7.3 for early Tega design sketches).

First, some background on our stakeholders and our co-design methods. The stakeholders with whom we engaged included

Figure 7.3
Concept sketches from the early design phase of Tega.

parents, teachers, school administrators, early childhood development experts, and children from Boston-area public schools that serve households from a variety of socioeconomic backgrounds. We made special efforts to include stakeholders from ethnically and linguistically diverse backgrounds, including bilingual and English-language-learning children and families. We used a variety of methods, including meetings, surveys, interviews, and focus groups, to learn about stakeholders' values and perspectives on using relational robots in early childhood education. Our co-design methods were iterative: we would have discussions with stakeholders, go back to our lab to integrate their perspectives and values into our design work, come back to the stakeholders for more discussion and feedback, and so on.

We also developed co-design methods specifically aimed at children. We brought children and parents *together* into the lab to interact with lower-fidelity relational robot prototypes (i.e., prototypes that did not include all the features we might deploy in a robot in a school). These were often remotely controlled by a person (as opposed to being autonomous)—this method is known as *Wizard of Oz*.[7] This prototyping method helped us understand the types of emotional interactions children would want to have with a robot, and crucially, it helped us do so *before* we built any AI algorithms that powered child–robot interactions completely autonomously. We also developed simple games and picture-based questionnaires for children, like the sorting activity discussed in the subsection "Inauthenticity as Unreality?" In one questionnaire, we asked children about their perceptions of Tega's social and relational attributes (e.g., "Let's pretend Tega didn't have any friends. Would Tega not mind or would Tega feel sad?" and "Does Tega really like you, or is Tega just pretending?"); children could point to pictures of Tega in their responses as well as explain their thinking. We invited children to draw pictures to different prompts, including many about potential relationships they might form with Tega, such as, "Draw a picture of your dream robot and what you do together."

Co-design approaches had material impacts on how we built robots like Tega. In our early discussions with stakeholders, we identified a widely held assumption: parents and teachers frequently assumed that robots like Tega would take the role of a teacher—that is, that Tega would relate to children as a source of authority and expertise. (This is not surprising given that many research labs and companies are developing intelligent expert tutoring systems, like Squirrel Ai and COLit.[8]) However, when we talked with parents, teachers, and children about how they *wanted* Tega to relate to children, we heard a different message. Many believed that children's educational needs would be better served if relational robots were to take the role of a *peer-like learning companion* as opposed to an expert teacher (e.g., Chen et al., 2020).

Stakeholders offered a variety of reasons for preferring a peer-like robot over a teacher-like robot. Teachers explained that they saw value in a robot that could be used as a "motivator or reinforcer," provide a "non-judgmental safe learning space," and introduce children to "activities they might not otherwise do" (Kory-Westlund et al., 2016)—all things they believed would be more easily achieved with a robot in a peer-like role. Teachers also expressed concerns that if the robot were to take on a teacher-like role, children would perceive it as competing with human teachers in the classroom. Further, teachers worried that a teacher-like robot might be more likely to "replace" teachers in the future; this, teachers believed, would harm how children learned and could result in teachers losing their jobs. Children, too, expressed a preference for engaging in peer-like relationships with robots. They responded more positively to a robot that asked them to play as another child would ("Do you want to play a story game?") than a robot that directed the activity in a teacher-like way ("Let's practice our storytelling now"). Children also reacted positively and learned more effectively when robots appeared friendly and inviting, like a special kind of pet rather than a distant authority figure. Children favored plush fabrics and bright, contrasting colors, often petting the robot or putting their arm around it as they played games together (see figure 7.4). These preferences are also in line with existing research suggesting that peer-based learning improves educational outcomes and brings motivational, cognitive, social, and emotional benefits for peers involved (Damon, 1984; Hassinger-Das et al., 2016; Topping, 2005 Tudge, 1989).

In light of these preferences, we adjusted our designs: rather than designing the robot as an expert teacher, we cultivated a child–robot relationship by designing Tega to be a peer-like or pet-like learning companion. For example, we programmed Tega to use language that is more child-like (and less teacher-like), such as the language mentioned in the previous paragraph. We also designed

Figure 7.4
A child with Green the DragonBot.

Tega to occasionally make mistakes—for example, Tega sometimes incorrectly answers questions about vocabulary or the content of a story—to make it appear less authoritative (and to allow it to model a growth mindset; see the subsection “Why Build Relational Robots for Early Childhood Education?”). Based on children’s interactions and preferences, we chose bright, soft material and a cute, animal-like design so that the robot would look like a kind of special, friendly pet.

These design choices had the intended effect—children in our studies tended to relate to the robot as a pet or playmate (Kory-Westlund, 2019; Kory-Westlund & Breazeal, 2019a, 2019b; Kory-Westlund et al., 2018; Park et al., 2019). They assumed the robot liked playing with them, too: “I know Green [the robot] likes to play with me, so I know he’s happy!” (Kory, 2014).

When we invited parents and guardians to participate in co-design sessions, we made further discoveries about what kinds of child–robot relationships we should design. We learned that many parents wanted to be involved as their children learned with Tega. We thus designed Tega to engender a *group relationship* among children, robots, and adults. For example, we created a special French-language-learning activity for Tega and asked 16 families to participate so we could hear their feedback and perspectives. As part of the activity, the robots used only French words when conversing with children. Parents participated in the learning activity by pointing out (in English) to the child when the robot was using new words and then prompting the child to repeat or use that word: “How do you say ‘bye’ in French?” (Freed, 2012). The robot facilitated French learning by indirectly prompting the parent to guide and teach their child. Parents told us that they experienced a *socially inclusive* experience, contrasting it with what they saw as socially exclusive experiences they have when their child uses a tablet (like an iPad). It would not have been possible to understand the importance and value of these group relationships without the close collaboration with parents and guardians as co-designers.

Conclusion

In an interview in the *Guardian*, Sherry Turkle warned that "if people start to buy the idea that machines are great companions . . . , as they increasingly seem to do, we are really playing with fire" (Adams, 2015). We agree with Turkle that developing relational robots raises genuine social and ethical concerns. But we also believe that, when designed and implemented responsibly, these technologies have the potential to serve as tools for helping to achieve transformative change. We've argued that to responsibly build relationships between children and robots, and to address concerns about authenticity, co-design is required. Stakeholders deserve a say in deciding what kinds of child–robot relationships (if any) we should design. If we want to "avoid playing with fire," all of us need to be in this together.

Notes

1. See, for example, Berscheid & Reis (1998); Csikszentmihalyi & Halton (1981); Kelley et al. (1983).

2. For more on Buddy, see http://www.bluefrogrobotics.com/; on Jibo, see https://www.jibo.com/; on Mabu, see http://www.cataliahealth.com/; on Alexa, see Sciuto et al. (2018). For academic work on Aibo, see, e.g., Fink et al. (2012); Friedman et al. (2003); Kahn et al. (2002); Weiss et al. (2009).

3. For a representative sample of work, see Breazeal et al. (2016); Chen et al. (2020); Kory-Westlund (2019); Kory-Westlund & Breazeal (2019a, 2019b); Westlund et al. (2017); Park et al. (2017).

4. Other researchers and scholars have also weighed in on the question of authenticity, e.g., Coeckelbergh (2012); Picard & Klein (2002). See also additional work by Turkle (2005, 2017).

5. We're using *connection* as a general term that encompasses relationships.

6. We don't mean to suggest that co-design is the only appropriate or useful methodology for the responsible design of relational robots. The responsible design of

any technology requires many complementary approaches, including those related to legal compliance, monitoring and assessment, and data governance. For details of other approaches, see, e.g.: the Montreal Declaration for Responsible AI (n.d.); the Institute of Electrical and Electronics Engineers' recommendations on ethically aligned design (IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019); the European Group on Ethics in Science and New Technologies statement on artificial intelligence, robotics, and "autonomous" systems (European Group on Ethics in Science and New Technologies, 2018); and value-sensitive design (e.g., Friedman & Hendry, 2019).

7. Wizard of Oz is a common technique enabling researchers to explore aspects of interaction not yet backed by autonomous systems. See Riek (2012).

8. For more details on these systems, see squirrelai.com/; Cole et al. (2007); Wise et al. (2005).

References

Adams, T. (2015, October 18). Sherry Turkle: 'I am not anti-technology, I am pro-conversation.' *The Guardian*. http://www.theguardian.com/science/2015/oct/18/sherry-turkle-not-anti-technology-pro-conversation

Ali, S., Moroso, T., & Breazeal, C. (2019). Can children learn creativity from a social robot? In *Proceedings of the 2019 on Creativity and Cognition (C&C '19)* (pp. 359–368). Association for Computing Machinery. https://doi.org/10.1145/3325480.3325499

Ames, M. G. (2019). *The charisma machine: The life, death, and legacy of one laptop per child*. The MIT Press.

And Also Too. (n.d.). *What is co-design?* Retrieved December 8, 2020, from http://www.andalsotoo.net/what-is-codesign

Arnold, L., Lee, K. J., & Yip, J.C. (2016). *Co-designing with children: An approach to social robot design* [paper presentation]. The 11th AMC/IEEE International Conference on Human–Robot Interaction, Christchurch, New Zealand, March 7–10. http://bigyipper.com/wp-content/uploads/2012/05/Co-designing-with-children-An-approach-to-social-robot-design.pdf

Banta, D. (2009). What is technology assessment? *International Journal of Technology Assessment in Health Care*, *25*(S1), 7–9.

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vols. 1–2, pp. 193–281). McGraw-Hill.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *12*(2), 293–327.

Breazeal, C., Harris, P. L., DeSteno, D., Kory-Westlund, J. M., Dickens, L., & Jeong, S. (2016). Young children treat robots as informants. *Topics in Cognitive Science*, *8*(2), 481–491. https://doi.org/10.1111/tops.12192

Boteanu, A., Chernova, S., Nunez, D., & Breazeal, C. (2016). Fostering parent–child dialog through automated discussion suggestions. *User Modeling and User-Adapted Interaction*, *26*(5), 393–423.

Chang, A., Breazeal, C., Faridi, F., Roberts, T., Davenport, G., Lieberman, H., & Montfort, N. (2012). Textual tinkerability: Encouraging storytelling behaviors to foster emergent literacy. *CHI '12 Extended Abstracts on Human Factors in Computing Systems* [Conference proceedings], 505–520.

Chen, H., Park, H. W., & Breazeal, C. (2020). Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Computers & Education*, *150*, 103836.

Coeckelbergh, M. (2012). Are emotional robots deceptive? *IEEE Transactions on Affective Computing*, *3*(4), 388–393.

Cole, R., Wise, B., & van Vuuren, S. (2007). How Marni teaches children to read [Special issue: Pedagogical Agents]. *Educational Technology*, *47*(1), 14–18.

Costanza-Chock, S. (2020). *Design justice*. The MIT Press.

Csikszentmihalyi, M., & Halton, E. (1981). *The meaning of things: Domestic symbols and the self*. Cambridge University Press.

Damon, W. (1984). Peer education: The untapped potential. *Journal of Applied Developmental Psychology*, *5*, 331–343.

Druin, A. (2002). The role of children in the design of new technology. *Behaviour & Information Technology*, *21*, 1–25.

Dweck, C. S. (2008). *Mindset: The new psychology of success*. Ballantine Books.

European Group on Ethics in Science and New Technologies. (2018). *Statement on artificial intelligence, robotics and 'autonomous' systems*. Retrieved December 8, 2020, from https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source-78120382

Fink, J., Mubin, O., Kaplan, F., & Dillenbourg, P. (2012). Anthropomorphic language in online forums about Roomba, AIBO and the iPad. *Advanced Robotics and Its Social Impacts (ARSO), 2012 IEEE Workshop On*, 54–59.

Freed, N. A. (2012). *"This is the fluffy robot that only speaks French": Language use between preschoolers, their families, and a social robot while sharing virtual toys* [Master's thesis, Massachusetts Institute of Technology]. DSpace@MIT.

Friedman, B., & Hendry, D. (2019). *Value sensitive design: Shaping technology with moral imagination*. The MIT Press.

Friedman, B., Kahn, P. H., Jr., & Hagman, J. (2003). Hardware companions? What online AIBO discussion forums reveal about the human-robotic relationship. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 273–280.

Garcia, E., & Weiss, E. (2017). *Education inequalities at the school starting gate: Gaps, trends, and strategies to address them*. [Report]. Economic Policy Institute.

Gooberman-Hill, R., Horwood, J., & Calnan, M. (2008). Citizens' juries in planning research priorities: Process, engagement and outcome. *Health Expectations*, *11*(3), 272–281. https://doi.org/10.1111/j.1369-7625.2008.00502.x

Gordon, G., Breazeal, C., & Engel, S. (2015). Can children catch curiosity from a social robot? *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 91–98.

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes Publishing.

Hassinger-Das, B., Ridge, K., Parker, A., Golinkoff, R., Hirsh-Pasek, K., & Dickinson, D. (2016). Building vocabulary knowledge in preschoolers through shared book reading and gameplay. *Mind, Brain, and Education*, *10*(2), 71–80.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the High/Scope Perry Preschool Program. *Journal of Public Economics*, *94*(1–2), 114–128.

Hennen, L. (2012). Why do we still need participatory technology assessment? *Poiesis & Praxis*, *9*, 27–41. https://doi.org/10.1007/s10202-012-0122-5

Huijnen, C. A. G. J., Lexis, M. A. S., Jansens, R., & de Witte, L. P. (2017). How to implement robots in interventions for children with autism? A co-creation study involving people with autism, parents and professionals. *Journal of Autism and Developmental Disorders*, *47*(10), 3079–3096.

Hurlbut, J. B., Jasanoff, S., Saha, K., Ahmed, A., Appiah, A., Bartholet, E., Baylis, F., Bennett, G., Church, G., Cohen, I. G., et al. (2018). Building capacity for a global genome editing observatory: Conceptual challenges. *Trends in Biotechnology*, *36*(7), 639–641.

IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems* (1st ed.). Retrieved December 8, 2020, from https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html

Kahn, P. H., Friedman, B., & Hagman, J. (2002). I care about him as a pal: Conceptions of robotic pets in online AIBO discussion forums. *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, 632–633.

Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human robot interaction. *IEEE Transactions on Robotics*, *23*(5), 962–971.

Kelley, H. H., Berscheid, E., Christensen, A., Harvey, J., Huston, T., Levinger, G., McClintock, E., Peplau, L., & Peterson, D. (1983). *Close relationships*. Freeman.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*(4), 664–687.

Kory, J. (2014). *Storytelling with robots: Effects of robot language level on children's language learning* [Master's thesis, Massachusetts Institute of Technology]. DSpace@MIT.

Kory-Westlund, J. M. (2019). *Relational AI: Creating long-term interpersonal interaction, rapport, and relationships with social robots* [Doctoral thesis, Massachusetts Institute of Technology]. DSpace@MIT.

Kory-Westlund, J. M., & Breazeal, C. (2019a). A long-term study of young children's rapport, social emulation, and language learning with a peer-like robot playmate in preschool. *Frontiers in Robotics and AI*, *6*. https://doi.org/10.3389/frobt.2019.00081

Kory-Westlund, J. M., & Breazeal, C. (2019b). Exploring the effects of a social robot's speech entrainment and backstory on young children's emotion, rapport, relationship, and learning. *Frontiers in Robotics and AI*, *6*. https://doi.org/10.3389/frobt.2019.00054

Kory-Westlund, J. M., & Breazeal, C. (2019c). Assessing children's perceptions and acceptance of a social robot. *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, 38–50.

Kory-Westlund, J. M., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction*, *13*, 1–9.

Kory-Westlund, J. M., Gordon, G., Spaulding, S., Lee, J. J., Plummer, L., Martinez, M., Das, M., & Breazeal, C. (2016). Lessons from teachers on performing HRI studies with young children in schools. *The Eleventh ACM/IEEE International Conference on Human-Robot Interaction*, 383–390.

Kory-Westlund, J. M., Jeong, S., Park, H. W., Ronfard, S., Adhikari, A., Harris, P. L., DeSteno, D., & Breazeal, C. (2017). Flat versus expressive storytelling: Young children's learning and retention of a social robot's narrative. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00295

Kory-Westlund, J. M., Park, H. W., Williams, R., & Breazeal, C. (2018). Measuring young children's long-term relationships with social robots. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 207–218.

McBride, N. (2020). Robot enhanced therapy for autistic children: An ethical analysis. *IEEE Technology and Society Magazine*, *39*(1), 51–60. https://doi.org/10.1109/MTS.2020.2967493

Moharana, S., Panduro, A. E., Lee, H. R., & Riek, L. D. (2019). Robots for joy, robots for sorrow: Community based robot design for dementia caregivers. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 458–467.

Montreal Declaration Responsible AI. (n.d.). *The Declaration*. Retrieved December 8, 2020, from https://www.montrealdeclaration-responsibleai.com/the-declaration

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism* (1st ed.). PublicAffairs.

Nuñez, D. S. (2015). *GlobalLit: A platform for collecting, analyzing, and reacting to children's usage data on tablet computers* [Master's thesis, Massachusetts Institute of Technology]. DSpace@MIT.

Pallikkathayil, J. (2019). The truth about deception. *Philosophy and Phenomenological Research*, *98*(1), 147–166.

Park, H. W., Grover, I., Spaulding, S., Gomez, L., & Breazeal, C. (2019, July). A model-free affective reinforcement learning approach to personalization of an autonomous social robot companion for early literacy education. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(1), 687–694.

Park, H. W., & Howard, A. M. (2015). Retrieving experience: Interactive instance-based learning methods for building robot companions. *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 6140–6145.

Park, H. W., Rosenberg-Kima, R., Rosenberg, M., Gordon, G., & Breazeal, C. (2017). Growing growth mindset with a social robot peer. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 137–145.

Picard, R. W., & Klein, J. (2002). Computers that recognise and respond to user emotion: Theoretical and practical implications. *Interacting with Computers*, *14*(2), 141–169.

Riek, L. D. (2012). Wizard of Oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, *1*(1), 119–136.

Šabanović, Selma. (2010). Robots in society, society in robots. *International Journal of Social Robotics*, *2*, 439–450.

Sample, M., Boulicault, M., Allen, C., Bashir, R., Hyun, I., Levis, M., Lowenthal, C., Mertz, D., Montserrat, N., Palmer, M. J., et al. (2019). Multi-cellular engineered living systems: Building a community around responsible research on emergence. *Biofabrication, 11*(4), 1–9. https://doi.org/10.1088/1758-5090/ab268c

Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018). "Hey Alexa, what's up?": A mixed-methods studies of in-home conversational agent usage. *Proceedings of the 2018 Designing Interactive Systems Conference*, 857–868.

Serholt, S., & Barendregt, W. (2016). Robots tutoring children: Longitudinal evaluation of social engagement in child–robot interaction. *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, Article 64, 1–10. https://doi.org/10.1145/2971485.2971536

Singh, N. (2018). *Talking machines: Democratizing the design of voice-based agents for the home* [Master's thesis, Massachusetts Institute of Technology]. DSpace@MIT.

Sparrow, R., & Sparrow, L. (2006). In the hands of machines? The future of aged care. *Minds and Machines*, *16*, 141–161.

Street, J., Duszynski, K., Krawczyka, S., Braunack-Mayera, A. (2014). The use of citizens' juries in health policy decision-making: A systematic review. *Social Science & Medicine*, *109*, 1–9. https://doi.org/10.1016/j.socscimed.2014.03.005

Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, *25*(6), 631–645. https://doi.org/10.1080/01443410500345172

Torgesen, J. K. (2004). Preventing early reading failure—and its devastating downward spiral. *American Educator*, *28*(3), 6–19.

Tudge, J., & Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In M. H. Bornstein & J. S. Bruner (Eds.), *Interaction in human development* (pp. 17–40). Lawrence Erlbaum Associates.

Turkle, S. (2005). *The second self: Computers and the human spirit*. The MIT Press.

Turkle, S. (2007). Authenticity in the age of digital companions. *Interaction Studies*, *8*(3), 501–517.

Turkle, S. (2017). *Alone together: Why we expect more from technology and less from each other*. Hachette UK.

Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican day care. *Developmental Psychology*, *28*(6), 1106–1114. https://doi.org/10.1037/0012-1649.28.6.1106

Watts Belser, J. (2016). Vital wheels: Disability, relationality, and the queer animacy of vibrant things. *Hypatia*, *31*(1), 5–21.

Weiss, A., Wurhofer, D., & Tscheligi, M. (2009). "I love this dog"—Children's emotional attachment to the robotic dog AIBO. *International Journal of Social Robotics*, *1*(3), 243–248.

White, P. Q. (2021). Honesty and discretion. *Philosophy and Public Affairs*, *50*(1), 6–49.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*(4), 552–559. https://doi.org/10.1037/0012-1649.24.4.552

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

Wise, B., Cole, R., van Vuuren, S., Schwartz, S., Snyder, L., Ngampatipatpong, N., Tuantranont, J., & Pellom, B. (2005). Learning to read with a virtual tutor: Foundations to literacy. In C. Kinzer and L. Verhoeven (Eds.), *Interactive literacy education: Facilitating literacy environments through technology* (1st ed., pp. 31–75).

8

Early Adolescents' Perspectives on Digital Privacy

Nicholas D. Santer, Adriana Manago, Allison Starks, and Stephanie M. Reich

A vast array of information communication technologies (ICTs) permeates public-private boundaries at home and school (Livingstone, 2005; Taylor & Rooney, 2016), creating a perfect storm in early adolescence, when burgeoning needs for autonomy, exploration, self-expression, and peer connectedness make youth easy targets for "dataveillance" (Smith & Shade, 2018). *Dataveillance* refers to exploiting or commercializing children's play or social networking data as a capitalist commodity. In the US, the Federal Trade Commission restricts internet companies from collecting personal information on children younger than 13 through the 1998 Children's Online Privacy Protection Act (COPPA) based on the belief that children and young adolescents need greater protection because they do not have the same knowledge, experience, and self-regulatory capacities as adults to fully consent to and understand the consequences of their interactions with ICTs (Costello et al., 2016). Interventions to protect children and young adolescents should be closely informed by the knowledge, perspectives, and tendencies of youth themselves, yet most psychological research on privacy in digital media has focused on adults, older adolescents, or emerging adults (e.g., Agosto & Abbas, 2017; Wisniewski, 2018).

Thus, we know little about perspectives on privacy during a sensitive period of identity and autonomy development when youth begin to use social media (Shin et al., 2012; Shin & Kang, 2016; Youn, 2009).

To fill this gap, we surveyed young adolescents in two US public middle schools about their privacy knowledge, preferences, and practices. The survey asked students about their beliefs and attitudes toward prescriptive messages about privacy (e.g., apps are selling your information to advertisers, digital ink is permanent), their preferences in negotiating privacy trade-offs (e.g., prioritizing self-disclosure for peer connection and convenience over protection of personal information), and their privacy protection behaviors (e.g., turning location sharing off, keeping passwords from others). Our goal was to shed light on how young adolescents view digital privacy and how their privacy behaviors may be shaped by circumstances unique to this developmental period.

Why Is Privacy an Issue with Young Adolescents?

Digital technologies blur public-private boundaries of the past, making data from within the home more accessible to corporations and allowing for easier transfer and collection of data about children's lives in novel ways that are challenging for families to regulate (Smith & Shade, 2018; Taylor & Rooney, 2016). Nowadays, children are wirelessly connected to others and to databases from a very young age through smart toys, in-home personal assistants (e.g., Alexa, Google Assistant), and games and apps (e.g., YouTube Kids) (Holloway & Green, 2016). As youth transition from childhood to adolescence, they are increasingly compelled to communicate with their peers using social media such as Snapchat or Instagram. The challenge of regulating the rapid proliferation of social media marketed to younger adolescents (e.g., TikTok, Amino, Discord, Whisper, Kik) and the lack of legitimate protections baked into the Internet of Things (IoT) create new privacy risks for youth growing up in the digital age (Smith & Shade, 2018). During the COVID-19 pandemic, the transition to remote or hybrid learning also meant that many children and their families were compelled to use e-learning software on personal or home devices, where privacy settings vary widely, and where data sharing practices remain opaque for families to understand (Harwell, 2020; Maalsen & Dowling, 2020; Teräs et al., 2020).

One concern about privacy stems from *datafication* practices. Because children are being treated as algorithmic assemblages from an early age, their "complexities, potentialities, and opportunities are becoming restricted" (Lupton & Williamson, 2017, p. 787) through multifaceted corporate surveillance. The term *algorithmic assemblages* refers to reducing children's online activities into data points so that algorithms can better predict purchasing or viewing behaviors. In short, it describes how corporations develop datafied representations of children. As Haggerty and Ericson (2000) describe, these algorithmic assemblages occur when data representations of children are constructed from "a series of discrete flows" and "reassembled" to represent a person (p. 605) in ways that profit corporations and data analytics firms. Children come to be treated as combinations of data points because of the datafication and surveillance practices baked into the IoT toys, smart devices, and social media they use, which reduces their digital actions to abstract demographic information and preferences for a litany of products and online services (Rooney, 2012; Smith & Shade, 2018). As children and young adolescents increasingly engage with content on an ever-expanding variety of devices, media industries are iterating on ways to predict their behaviors and target them with personalized marketing messages based on these algorithmic reassemblages (Marx & Steeves, 2010; Regan & Steeves, 2019). Zuboff (2019) argues that human autonomy will be diminished under what she calls “surveillance capitalism,” where corporations exploit private experience for profit and trade on human behavioral futures.

Although notable legislation such as COPPA in the US is meant to deter children under 13 from creating different forms of social media accounts (e.g., Facebook, Instagram, YouTube), it is usually ineffective. There is no evidence that age limits work (children can lie about their age online), and their data are often still collected via parentally controlled accounts or apps (Federal Trade Commission, 1998; Smith & Shade, 2018). Moreover, social media use in the US tends to begin in middle school (Anderson & Jiang, 2018; Rideout & Robb, 2019), yet after age 13, adolescents are no longer protected under COPPA, leaving them potentially vulnerable in an essentially unregulated, commercial, digital media environment. COPPA’s age cutoff assumes that adolescents have the knowledge and maturity to act in their own interests regarding privacy on social media, but we know little about whether that assumption is warranted (Costello et al., 2016).

In early adolescence, youth are still developing both “cold” psychosocial systems of information processing (e.g., logical reasoning, analytical skills) and “hot” psychosocial systems of information processing (e.g., emotional, impulsive, experiential) (Costello et al., 2016). Steinberg et al. (2009) showed that there were significant differences in how young and older adolescents handled decision-making tasks related to psychosocial maturity (e.g., tasks involving risk perception, sensation seeking, impulsivity, future orientation) but not for tasks that tested their cognitive capacities (e.g., logical reasoning ability, information processing). The young adolescent brain is more susceptible to social and emotional factors (e.g., peer pressure, romantic attachment), and their capacity for regulating their behavior is still incomplete (Albert & Steinberg, 2011; Steinberg et al., 2009). Even older adolescents may still show social and emotional maturity deficits compared to adults (Cohen et al., 2016; Costello et al., 2016), especially for decision-making and reasoning about *incentives* (Galvan et al., 2006; Somerville et al., 2011), *threats/risks* (Rudolph et al., 2017), and *peers* (Chein et al., 2011).

The psychosocial factors (e.g., impulsivity, peer pressure, future orientation) that influence judgment and decision-making in early adolescence may impact the time it takes youth to activate "cold" analytical decision-making skills required for instantaneous digital privacy decisions (e.g., posting a sexy selfie, sharing photos of drug and alcohol use, revealing a location). Young adolescents are more likely to be motivated by "hot" experiential desires such as social connectedness, peer pressure, and self-presentation/identity presentation compared to older adolescents (Juvonen, 2007; Juvonen & Murdock, 1995; Knifsend & Juvonen, 2013). They are also less able to accurately weigh the risks versus rewards of using social media given these social and emotional pressures as well as the immediate gratification of quantified social feedback (e.g., Facebook or Instagram likes, Snapstreaks). Therefore, young adolescents may require additional protections (e.g., longer delays for reflection before posting content, for waiving privacy protections, and for sharing location) to continue developing their cognitive-control skills on social media in a way that does not compromise their needs for self-expression and identity development. These protections could also help bridge the gap of COPPA after age 13.

Technopanics and alarmist narratives transmitted through parents, teachers, and popular media have not been helpful in dealing with this issue because they tend to pathologize adolescents' use of social media (Agosto & Abbas, 2017; Marwick, 2008). Livingstone (2008, 2014) argues that technopanics can result in adolescents learning prescriptive messages (e.g., don't talk to strangers online, don't disclose personal information) without changing their behavior, partly because they do not see social media in the same terms as adults in their lives do—their main goals are not generally to meet strangers or disclose intimate personal information but to expand their social networks and build relationships

(Livingstone, 2008, 2014). Much of the research focusing on social media safety for youth is grounded in adults' prescriptive views of youths' attitudes and behaviors—what adults think youth should be doing online, as opposed to an informed view of what youth are actually doing online and why. As a result, adolescents' use of social media is framed as poor and risky (Livingstone, 2008; Marwick & boyd, 2014), leading to solutions narrowly aimed at reducing teens' online disclosures (Wisniewski, 2018). To address the lack of youth perspectives on digital privacy, we designed a survey to tap into privacy knowledge, preferences, and practices that may be unique to early adolescence.

What Are Adolescents' Perspectives on Privacy?

A developmental lens is necessary for understanding young adolescents' perspectives and behaviors related to digital privacy. Developmental needs for peer intimacy, connection, and identity exploration and formation introduce unique circumstances during this period of the life span. Evidence suggests that although adolescents can understand and reason about some risks like adults do, they are more sensitive to peer social rewards in risk-taking scenarios (Albert et al., 2013; Smith et al., 2014). Younger adolescents are more concerned with social connectedness and superficial self-presentation/identity presentation than older teens and adults (Juvonen, 2007; Juvonen & Murdock, 1995; Knifsend & Juvonen, 2013), which could lead to relinquishing more security in certain arenas to gain social validation and belonging—for example, disclosing personal information publicly to participate in online communities and accrue many likes, comments, and followers (Yau & Reich, 2019). In short, adolescents may be dealing with privacy trade-offs differently from adults as they negotiate incentives particular to this developmental period.

Proximal versus distal privacy is likely to be an important distinction in adolescents' perspectives on privacy. Adolescents' proximal, person-to-peer privacy management involves needs for intimacy, affiliation, exploration, and information control in everyday relationships, peer groups, and families (Peter & Valkenburg, 2011; Robinson, 2016). For example, adolescents' attitudes about online privacy and safety are often shaped by their discomfort with unintended audiences seeing their personal information, yet most youth feel pressure to share their personal information with friends to stay socially connected (Agosto & Abbas, 2017; Shin & Kang, 2016). This pressure reflects their desires to "be *in* public without always *being* public" (Marwick & boyd, 2014, p. 1052). Young adolescents must learn to balance their desires for social connectivity with their desires to restrict personal information from unintended audiences (e.g., parents, teachers, predators).

Whereas adolescents' proximal privacy management is influenced by their needs for identity development and social connectedness, distal privacy management involves developing a more complex understanding of how corporations and data brokers collect and trade on personal information. Adolescents' person-to-corporation privacy management could be challenging for tweens who are just beginning to think abstractly and understand issues at the level of society. For example, Shin et al. (2012) found that 9–12-year-olds tend to overestimate their understanding and invulnerability, perceiving themselves as more competent and knowledgeable in using ICTs than their parents but more willing to disclose personal information for marketing purposes. Adolescents could also be less disturbed by abstract invasions of privacy from governments, criminals, or corporations compared to the more immediate risks of nosy parents or peers (Marwick & boyd, 2014; Tufekci, 2008). Research with older adolescents and emerging adults found that they expressed little concern about the future use of their personal data while also showing limited knowledge of the business practices involved in using such information for commercial purposes (Lapenta & Jørgensen, 2015; Montgomery et al., 2017).

In our study, we distinguished between proximal and distal privacy by separately examining adolescents' perspectives on privacy regarding their social relationships and their perspectives on privacy regarding corporations and potential criminals. That is, we asked adolescents about their beliefs, attitudes, and behaviors related to protecting themselves in social networks, from corporate surveillance, and from predators. We also differentiated adolescents' preferences in their digital privacy negotiations in terms of their willingness to trade off some security for rewards via corporate surveillance (e.g., trading personal information for convenience) versus peer relationships (e.g., trading personal information for more followers).

How Do Adolescents' Beliefs, Attitudes, and Preferences Translate to Behaviors?

Adolescents generally show less concern about privacy than adults (Gasser & Palfrey, 2008; Moscardelli & Liston-Heyes, 2004), especially in sharing personal information on social media (Jones et al., 2009). This may be because they have more know-how to protect their privacy. Miltgen and Peyrat-Guillard (2014) found a reverse privacy paradox in adolescence where lower privacy concerns were associated with greater use of protective strategies for personal data. In contrast, adults with higher privacy concerns practiced fewer privacy behaviors because they had less knowledge of online privacy strategies (Blank et al., 2014). Adolescents are "digital natives" (Baumann, 2010) and perhaps more self-confident internet users who take a greater degree of personal responsibility for managing their proximal online privacy, often due to their technological literacy (Livingstone, 2008; Wisniewski, 2018).

Indeed, throughout adolescence, youth increasingly express more positive attitudes toward data management and responsible social media use, and they become more confident in their ability to prevent privacy violations (Miltgen & Peyrat-Guillard, 2014). Although the likelihood of providing certain types of personal information online (e.g., photos of oneself or friends, school names, hometown, screen names from other social media) increases with age (Lenhart & Madden, 2007; Steeves & Webster, 2008), younger adolescents are more likely to restrict privacy settings for their social media profiles than older teens (Caverlee & Webb, 2008; Livingstone, 2008). Studies with adolescents aged 14–18 years have shown they often engage in privacy-protecting behaviors but mostly those aimed at broadly controlling how their information is revealed to others (proximal privacy) (Livingstone, 2006, 2008). For example, they use pseudonyms on social media (Miltgen & Peyrat-Guillard, 2014), restrict profile access (boyd & Hargittai, 2010), delete tags and photos from friends, and limit friend requests and social connections (boyd, 2014; Marwick & boyd, 2014).

Although adolescents may be competent in proximal privacy practices to manage their reputations and social relationships, they may be less attuned to distal privacy practices for restricting their information from corporations or criminals. Adolescents may display another privacy paradox in terms of how they resolve tension between their corporate/criminal privacy concerns and their motivations for self-presentation and social connectedness (Utz & Krämer, 2009). In other words, adolescents may report being concerned about corporations and criminals, but they are more motivated to risk digital disclosure for their higher priorities of peer belonging and identity validation. To contribute a more nuanced picture of how young adolescents' knowledge and preferences account for their privacy-protecting behaviors, we distinguished between privacy-protecting behaviors related to corporations and criminals (e.g., paying for apps to avoid ads, using strong passwords) and those related to social relations (e.g., letting a friend of a friend they have not met follow them on social media).

Current Study

We administered a survey to young adolescents, 11–14 years old, to answer three main research questions about their privacy knowledge, preferences, and behaviors in proximal and distal contexts.

Research question 1: What kinds of digital privacy-protecting behaviors do young adolescents report vis-a-vis social networks versus corporate surveillance and potential criminals?

Research question 2: What are young adolescents' beliefs, attitudes, and preferences for digital privacy vis-a-vis social networks versus corporate surveillance and potential criminals?

Research question 3: Do adolescents' beliefs, attitudes, and preferences predict their digital privacy-protecting behaviors after accounting for demographic characteristics such as age, race/ethnicity, and gender?

Demographic Considerations

Given that limited research finds differences in privacy behaviors between younger and older adolescents (e.g., Shin et al., 2012; Rideout & Robb, 2019) but has not broadly explored heterogeneity of practices within early adolescence, we consider how our findings vary across the middle school grades. Further, some research has found differences in privacy preferences between males and females (Tifferet, 2019; Youn & Hall, 2008). Thus, we consider how gender may relate to adolescents' preferences, attitudes, beliefs, and practices. Finally, research on online activities, surveillance, and privacy often finds different beliefs and practices between White users and people of color (Madden, 2017; Shelley et al., 2004). We suspect that Whiteness and its corresponding systemic privilege will be demonstrated by White students being less concerned about corporate surveillance while practicing greater proximal privacy-protecting behaviors than students of color. However, given the heterogeneity of samples from students who are not White, we do not have hypotheses about attitudes, beliefs, and behaviors between different racial or ethnic groups.

Method

Youth aged 11–14 years in the Southeastern United States were asked to complete an online privacy survey during their regular school day. Two separate middle schools serving grades six through eight participated in the survey in May and June of 2019. Participating schools were part of a larger school district serving approximately 14,000 students from prekindergarten through grade 12 in a blend of rural, suburban, and urban settings covering 726 square miles. The school district used a one-to-one model of instructional technology integration, wherein students in the participating district were assigned a personal laptop to use throughout the academic school year. Surveys were administered through these laptops using Qualtrics.

At the time of the survey, Allison Starks (third author) was serving as a technology coach for the two participating middle schools. She collaborated with school leadership on multiple technology initiatives throughout the year, addressing a variety of instructional and technical needs related to technology integration. Specifically, school leaders were interested in digital literacy, including student privacy practices.

Participants Schools were selected based on convenience and interest in digital privacy habits. As described, two middle schools were offered a chance to participate. Both schools expressed interest in learning more about youth digital privacy behaviors in service of digital citizenship initiatives. Middle schoolers ($N = 414$) ages 11–14 completed the survey during class time. Participating students are described in table 8.1. They were 54 percent male ($n = 224$, and most identified as White (52.2 percent), multiracial/ethnic (12.3 percent), Latine (8.9 percent), Black (8.5 percent), or Asian/Asian American (8.0 percent). Seventy percent of youth reported using social media. Within age groups, 64 percent of students under 13 years old identified as social media users, compared with 76 percent of students 13 or older.

Measures The privacy survey was created by the authors, with assistance from central office administrators in the school district and school leadership at each site. School technology leaders (classroom teachers, principals, and librarians) also gave feedback that led us to revise the survey's content and length.

The anonymous survey asked youth about their gender, age, race/ethnicity, and the social media websites or apps they use. It asked about the frequency of privacy-protecting behaviors and the participants' attitudes and preferences around protecting privacy on social media, websites, and devices. Of the 14 behaviors asked, eight focused on explicit behaviors that protect against hacking or access by unknown others such as criminals or corporations and six asked about behaviors related to privacy in social relationships. Of the seven attitude questions, three focused on corporate surveillance, two on future-oriented thinking, and two on predators. Of the 10 items addressing digital privacy preferences, four questions asked about corporate surveillance, two about peer social relationships and four questions about peer relationship preferences that had potential risk. Table 8.3 contains sample questions and response formats for each of these subscales.

Because questions about privacy preferences were last on the survey, they had lower response rates due to time limits. Only 295 of the 414 students completed these privacy preference items before the end of class. Comparisons between those with and without

Table 8.1
Sample demographics—student characteristics

	N	%
Gender		
Male	224	54
Female	190	46
Age		
11	38	9.2
12	164	39.6
13	155	37.4
14	55	13.4
Nonreal value (e.g., 38)	2	0.4
Race/Ethnicity		
White	216	52.2
Multi racial/ethnic	51	12.3
Black or African American	35	8.5
Latino/a	37	8.9
Asian/Asian American	33	8.0
Native American or Pacific Islander	2	0.5
Prefer not to say	18	4.3
Other	22	5.3

preference data found no differences in age, gender, race/ethnicity, or social media activities. A comparison of demographics between the two groups can be found in table 8.2.

Total scores and subscale scores were calculated for each of the attitudes, beliefs, preferences, and practices described in table 8.3. To account for missing items, mean scores were calculated based on the number of items answered. For example, if a participant responded to only 10 of the 14 privacy-protecting behaviors, the mean number of behaviors was calculated with a denominator of 10 rather than 14. Although the analysis did not include a minimum number of answered items required for a scale score, all

Table 8.2
Comparison of demographics between analytic sample (n = 295) and those with incomplete data (excluded) (n = 119)

	Analytic sample	Incomplete/excluded sample
Female	45%	47%
White	50%	57%
Age		
11	9%	11%
12	39%	41%
13	37%	39%
14	15%	9%

students answered at least three items to have an average score included in the analyses. For interpretability, several items under privacy in social relationships were reverse coded (noted in table 8.3) so that higher scores indicate more privacy protection. Total privacy behaviors consisted of privacy behaviors across social relationships and in relation to corporations and criminals.

Analytic plan To address research questions 1 and 2, we describe the frequency of specific attitudes, preferences, and practices around privacy within social relationships, in relation to potential criminals such as hackers and phishers, and in connection to corporations or other business or data-brokering interests.

To address research question 3, we used ordinary least squares (OLS) regressions to determine if attitudes, beliefs, and preferences about privacy predicted privacy-protecting behaviors globally and by domain (e.g., beliefs that corporations are collecting information predicts behaviors that limit their ability to collect data). The regression models that included preferences as predictors used a smaller sample ($n = 295$), due to some students running out of time

Table 8.3
Privacy survey

Topic of questions	Specific items	Response options	Summary scores
Privacy-protecting behaviors			
Privacy against corporations and criminals 8 items	• Read privacy policies for apps/websites • Use strong passwords • Clear my browser history • Turn location sharing off • Use a VPN • Keep my social media accounts private • Block push notifications • Pay for apps to avoid ads	**Frequency** (1 = never, 2 = sometimes, 3 = most of the time, 4 = always)	Mean = 2.20 SD = 0.55 Range: 1–4
Privacy in social relationships 6 items	• Share personal information (R) • Let **only people I have met in person** follow me on social media • Let **a friend of a friend, who I have not met**, follow me on social media (R) • Let **people I do not know** follow me on social networking sites (R) • Ask people to take down photos of me/posts about me that I do not like • Share my passwords (R)	**Frequency** (1 = never, 2 = sometimes, 3 = most of the time, 4 = always)	Mean = 3.03 SD = 0.50 Range: 1–4

(*continued*)

Table 8.3
(*continued*)

Topic of questions	Specific items	Response options	Summary scores
Attitudes and beliefs about privacy			
Corporate surveillance 3 items	• Companies use information about what you do online to try and sell you things • Your phone is listening to what you say, so companies can target ads at you • Apps are selling your information to advertisers	**Likert** (1 = strongly disagree to 5 = strongly agree)	Mean = 3.08 SD = 1.05 Range: 1–5
Future-oriented thinking 2 items	• Having a positive reputation online is important for getting a job in the future • Digital ink is permanent (what is posted online never goes away)	**Likert** (1 = strongly disagree to 5 = strongly agree)	Mean = 3.53 SD = 1.00 Range: 1–5
Potential predators 2 items	• You can never be really sure who you are talking to online • It is important to keep all your accounts private	**Likert** (1 = strongly disagree to 5 = strongly agree)	Mean = 3.82 SD = 1.00 Range: 1–5

Table 8.3
(*continued*)

Topic of questions	Specific items	Response options	Summary scores
Preferences about privacy and sharing			
Peer social relationships 2 items	• I like it when a friend tags me in a positive post • Posting on social media is important for my friendships	**Likert** (1 = strongly disagree to 5 = strongly agree)	Mean = 3.00 SD = 0.98 Range: 1–5
Peer relationships with introduction of risk 4 items	• I like it when social media tags my location • I like it when people like my posts, even if I don't know them • It's okay to allow people to follow me, even if I don't know them • Having more followers makes me feel good about myself	**Likert** (1 = strongly disagree to 5 = strongly agree)	Mean = 2.71 SD = 0.93 Range: 1–5
Corporate surveillance 4 items	• I like it when apps suggest people I should follow or connect with • I like seeing ads for things I like online • I like it when I can log in with Google or Facebook • I like it when a website or app already knows who my friends are	**Likert** (1 = strongly disagree to 5 = strongly agree)	Mean = 2.91 SD = 0.91 Range: 1–5

Note: (R) indicates reverse coding; higher scores on subscales mean higher privacy protection.

during their advisory period. Including a dummy variable for missing student data in the regression analyses revealed no significant differences for outcomes of interest (e.g., total privacy behaviors, privacy-protecting behaviors in social relationships, and privacy-protecting behaviors against corporations and criminals). T-tests revealed no significant differences by age, gender, or race/ethnicity between the sample with preferences data ($n = 295$) and the sample with missing preferences data. To determine how demographic characteristics relate to privacy attitudes, beliefs, preferences, and practices, we included gender, age, and race/ethnicity as covariates in the research question 3 analyses.

For all analyses, grade, gender, and race/ethnicity were included as covariates. Grade was valued as sixth, seventh, or eighth grade. Gender was dichotomized as male or female since we were limited in asking about participants who identified as nonbinary based on schools' feedback on our survey instrument. We did add a third option to our gender question, "prefer not to say," but only 19 students selected this option. Given that most students were White and that Whiteness is likely associated with different privacy preferences, race/ethnicity was dichotomized as white or youth of color.

For survey data cleaning, we first distinguished between different types of incomplete nominal and ordinal data (i.e., "select all that apply" questions: privacy practices, preferences, beliefs or age, gender, race/ethnicity) in our initial sample ($n = 669$). Adolescents who had only completed demographics questions without finishing the rest of the survey were removed from our analyses, as were responses that contained gibberish or nonsensical content for self-report questions (e.g., social media used). Additionally, users who satisficed (e.g., straightlined) in their responses or those with extremely fast response times were also removed from our analyses. This left us with a final sample size of 414.

Results

RQ1: Behaviors To better understand young adolescents' privacy-protecting behaviors, we looked at the total number of behaviors, the frequency of these behaviors, and specific behaviors promoting privacy for social relationships and with corporations and criminals. On average, youth engaged in nine discrete privacy behaviors at least some of the time ($M=8.73$, $SD=2.63$, $N=414$), ranging from three to 14 behaviors. Figure 8.1 details the percentage of youth engaging in each privacy behavior. Regarding patterns of behaviors, youth reported an average of five of eight possible privacy behaviors that protect against hacking or corporate interests ($M=4.71$, $SD=1.66$, range: 1–8) and an average of four of six possible behaviors that involve social relationships ($M=4.02$, $SD=1.42$, range: 1–6). In considering differences in privacy behaviors across different demographic characteristics, White youth reported significantly fewer total and specific privacy behaviors ($M=8.43$, $SD=2.68$, $n=216$) than their peers of color ($M=9.05$, $SD=2.54$, $n=198$; $t=2.39$, $p=0.02$). Specifically, White youth ($M=4.52$, $SD=1.66$) reported significantly fewer *privacy-protecting behaviors against corporations and criminals* than their peers of color ($M=4.92$, $SD=1.65$; $t(412)=2.43$, $p=0.02$). Youth also varied in the average number of privacy behaviors reported between age groups ($F(3, 408)=4.24$, $p<.01$), with younger students reporting fewer total privacy-protecting behaviors (11 years, $M=7.58$; 12 years, $M=8.50$; 13 years, $M=9.10$; 14 years, $M=9.01$). Youth did not vary in the number of *privacy against corporations or criminals* behaviors by age. However, younger students reported fewer *privacy behaviors in social relationships* compared to older students (11 years, $M=3.47$; 12 years, $M=3.87$, 13 years, $M=4.25$; 14 years, $M=4.18$; $F(3, 408)=4.16$, $p<.01$).

In terms of frequency, youth reported engaging in privacy-protecting behavior from *corporations and criminals* sometimes to

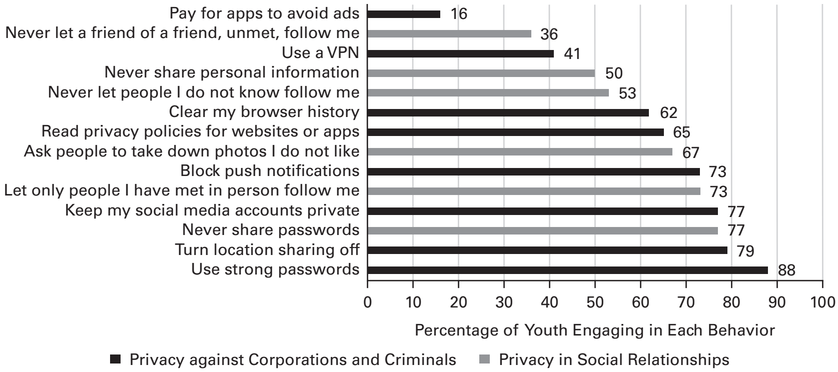


Figure 8.1
Youth engagement in privacy behavior.

most of the time ($M = 2.20$, $SD = 0.55$). Interestingly, youth more frequently engaged in behaviors that protect privacy in social relationships, reporting a mean of 3.03 ($SD = 0.5$), which equates to *most of the time* to *all of the time*, as shown in table 8.3. Average frequency of *privacy in social relationships* scores varied with age ($F(3, 408) = 2.81$, $p = 0.04$) with younger adolescents reporting more frequent privacy-protecting behaviors in social relationships than older children (11 years, $M = 3.17$; 12 years, $M = 3.07$, 13 years, $M = 3.00$, 14 years, $M = 2.91$).

In looking at specific types of behaviors, 56 percent of youth reported sharing their location with someone when using a smartphone. Most common location sharing partners were parents (22 percent), friends (19 percent), and siblings (17 percent). Interestingly, 79 percent reported turning off location sharing on apps. A total of 88 percent of youth report using strong passwords, while 77 percent reported never sharing their passwords with others.

RQ2: Beliefs, attitudes, and preferences To better understand adolescents' attitudes and beliefs about privacy (RQ2), we asked youth to rate their agreement with many well-known prescriptive messages surrounding digital privacy (e.g., You can never be really sure who you are talking to online, digital ink is permanent). These messages centered on *corporate surveillance*, *protection from predators*, and presentation online for *future oriented thinking*, as described in table 8.3. Frequencies of agreement with each prescriptive message are presented in figure 8.2.

Agreement with these prescriptive privacy messages was rather low. For instance, only 30 percent of youth believed that "Apps sell information to advertisers." This is especially interesting considering that 58 percent of youth agreed that "companies use information about what you do online to try and sell you things." The highest rate of agreement was with messages that focused on contact with strangers, which we labeled *protection against potential predators* (e.g., it is important to keep all your accounts private). Subscale scores for attitudes around corporate surveillance, future thinking, and protection from predators (see table 8.3) suggest that youth, on average, agree more frequently with messaging around *protection against predators* ($M=3.82$, $SD=1.00$) than messaging around *corporate surveillance* ($M=3.08$, $SD=1.05$) and *future-oriented thinking* ($M=3.53$, $SD=1.00$).

Attitudes about these messages differed across gender, with females agreeing more with future-oriented items ($M=3.72$, $SD=0.89$) than males ($M=3.38$, $SD=1.05$; $t=-3.47$, $p<.01$). White youth reported more agreement with *future-oriented* items ($M=3.66$, $SD=0.94$) than

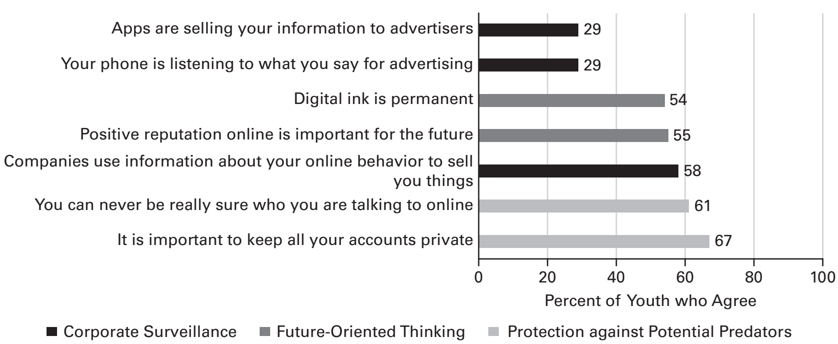


Figure 8.2
Youth agreement with prescriptive messages.

peers of color ($M = 3.40$, $SD = 1.03$; $t = -2.66$, $p < .01$). Privacy attitudes around *protection against predators* differed by gender as well, with females agreeing more with statements promoting privacy *protection against predators* ($M = 3.96$, $SD = 0.93$) compared with males ($M = 3.71$, $SD = 1.04$; $t = -2.51$, $p = .02$).

Of the 295 students who completed preference items, privacy-protecting preferences conceptually organized around exclusively peer social relationships (e.g., I like it when a friend tags me in a positive post), peer relationships with potential risks (e.g., I like it when people like my posts, even if I don't know them), and corporate surveillance (e.g., I like seeing ads for things I like online), as described in table 8.3. On average, youth were less likely to agree with preferences that *introduced risk into their social relationships* ($M = 2.71$, $SD = 0.93$; e.g., it's okay to allow people to follow me, even if I don't know them), where connecting with others may allow for connecting with potential predators, compared with situations facilitating *peer social relationships* ($F(21, 273) = 14.36$, $p < .001$; $M = 3.00$, $SD = 0.98$; e.g., posting on social media is important for my friendships). Additionally, youth were less likely to agree with preferences for *corporate surveillance* statements ($M = 2.91$, $SD = 0.91$; e.g., I like it when apps suggest people I should follow or connect with) compared with situations facilitating *peer social relationships* ($F(21, 273) = 8.07$, $p < .001$; $M = 3.00$, $SD = 0.98$). The frequency of youth digital privacy preferences is reported in figure 8.3, suggesting that popular youth preferences span categories of peer social relationships (e.g., I like it when a friend tags me in a positive post) and corporate surveillance (e.g., I like it when I can log in with Google or Facebook). Youths' preferences around corporations and peer relationships did not significantly differ by gender or race/ethnicity. However, females expressed less preference for privacy in peer relationships ($M = 2.58$, $SD = 0.88$) than males ($M = 2.81$, $SD = 0.96$; $t = 2.12$, $p = 0.04$).

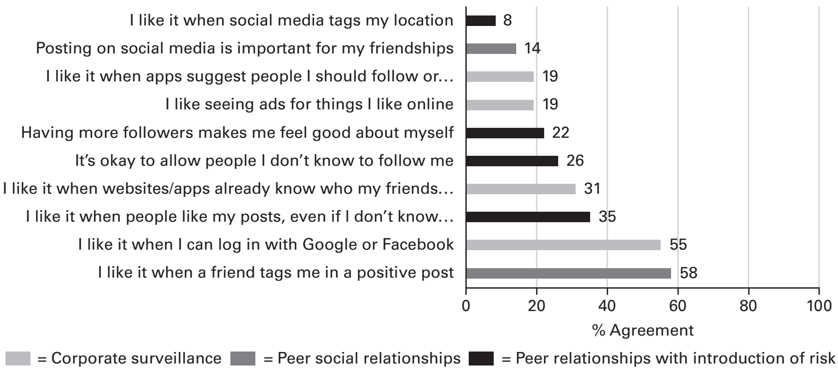


Figure 8.3
Frequency of youth digital privacy preferences.

RQ3: Predicting behaviors from attitudes, beliefs, and preferences Next, we considered how youths' beliefs, attitudes, and preferences, alongside sociodemographic factors, predicted their privacy-protecting behaviors.

Total privacy behaviors Privacy-protecting behaviors involve negotiating with social relationships, potential predators, and corporations. Using multivariate OLS regressions to predict total privacy-protecting behaviors based on beliefs/attitudes (Model 1) and preferences (Model 2) as shown in table 8.4, we found that youth with beliefs about protecting against predators online ($\beta = 0.13$, $p < .001$) were more likely to engage in more privacy-protecting behaviors in general, accounting for 13 percent of variance in behavior (Model 1). Including youths' preferences accounted for 27.4 percent of the variance in global privacy behaviors (Model 2). Attitudes around potential predators ($\beta = 0.15$, $p < .001$) (e.g., you can never be really sure who you are talking to online), less preference for peer relationships with potential risks ($\beta = -0.15$, $p < .001$) (e.g., I like it when people like my posts, even if I don't know them) (reversed scored), and more preference for corporate surveillance ($\beta = 0.07$, $p < .05$) (e.g., I like it when I can log in with Google or Facebook) were significant predictors of total privacy-protecting behaviors, holding other predictors constant.

Privacy in social relationships Youth engaged in different types of privacy behaviors based on whether privacy was for social relationships or to protect against corporations and potential criminals. Using adolescent attitudes around privacy and student demographics as predictors (Model 3), we found that youth with beliefs about protecting against predators online ($\beta=0.15$, $p<.001$) were more likely to engage in more privacy-protecting behaviors in social relationships, accounting for 15.6 percent of variance in behavior (Model 3). Incorporating preferences around privacy (Model 4), along with sociodemographic characteristics, increased the percentage of explained variance in privacy behaviors vis-a-vis social relationships (e.g., I let only people I have met in person follow me on social media) to 30 percent. Youths' beliefs and attitudes related to potential predators online ($\beta=0.11$, $p<.001$) and less preference for peer relationships with potential risk ($\beta=-0.22$, $p<.001$) (Model 4) were all significant predictors of privacy behaviors vis-a-vis social relationships, as shown in table 8.4.

Privacy against corporations and criminals Youths' privacy attitudes and student characteristics (Model 5) account for 8.8 percent of the variance in privacy against corporations and criminals, as shown in table 8.4. Incorporating youth preferences (Model 6) alongside attitudes and student characteristics accounted for 19 percent of the variance in behaviors related to privacy against corporations and criminals, as shown in table 8.4. Attitudes about protecting against potential predators predicted privacy behaviors that protect against predators, such as "use strong passwords" and "keep social media accounts private" ($\beta=0.18$, $p<.001$), while higher preference for peer relationships with risk was associated with less privacy-protecting behavior against corporations and criminals ($\beta=-0.09$ $p<.05$). Interestingly, greater preference for corporate surveillance (e.g., I like it

Table 8.4
Regression predicting global and specific privacy behaviors

	Global privacy behavior		Privacy in social relationships		Privacy against corporations and criminals	
	(1)	(2)	(3)	(4)	(5)	(6)
Age	.02 (.02)	.03 (.02)	–.08** (002)	–.08** (.03)	.10*** (.03)	.11*** (.03)
Female	.09* (.04)	.05 (.04)	.10* (.05)	.05 (.05)	.08 (.05)	.04 (.06)
White	–.03 (.04)	–.08 (.04)	.04 (.05)	–.02 (.05)	–.06 (.05)	–.12* (.06)
Corporate surveillance attitudes	.00 (.02)	–.00 (.02)	.00 (.03)	.00 (.03)	.00 (.03)	–.01 (.03)
Future-oriented thinking attitudes	.01 (.03)	.01 (.03)	.01 (.03)	.05 (.03)	–.00 (.03)	–.02 (.04)
Potential predators' attitudes	.13*** (.02)	.15*** (.03)	.15*** (.03)	.11*** (.03)	.13*** (.03)	.18*** (.04)
Corporate surveillance preferences		.07* (.03)		.04 (.04)		.09* (.04)
Peer relationship preferences		.02 (.03)		.01 (.04)		.03 (.04)
Peer relationships with risk preferences		–.15*** (.03)		–.22*** (.04)		–.09* (.04)
Constant	1.70*** (.28)	2.17*** (.33)	3.31*** (.33)	4.07*** (.40)	0.44 (.38)	.75 (.45)
Observations	401	293	401	293	401	293
R^2	.130	.274	.156	.300	.088	.190

Note: $^{*}p<.05$, $^{**}p<.01$, $^{***}p<.001$; standard error in parentheses. The outcome variables—global privacy behavior, privacy in social relationships, and privacy against corporations and criminals—are based on numbers of different types of behaviors endorsed.

when I can log in with Google or Facebook) was associated with more frequent privacy-protecting behavior against corporations and criminals ($\beta = 0.09$, $p < .05$).

Differences by student characteristics To understand the degree to which young adolescents differentiate in their privacy behaviors against corporations and criminals compared with privacy in social relationships, based on what we know about development, we used demographic characteristics as predictors for privacy behaviors globally and across domains. Specifically, age, gender, and race/ethnicity were used as covariates. Results, shown in table 8.4, indicate that age was a significant predictor of privacy behavior in social relationships and privacy behavior against corporations and criminals, though in different directions. Older students tended to engage in fewer privacy-protecting behaviors in social relationships ($\beta = -0.08$, $p < .01$) than younger adolescents. Alternately, behavior that protects against corporations and criminals was more frequent ($\beta = 0.11$, $p < .001$) among older tweens. Gender was not a significant predictor for privacy behaviors when controlling for attitudes and preferences, while race/ethnicity was related to behaviors against corporations and criminals. Notably, White youth ($\beta = -0.12$, $p < .05$) reported fewer privacy-protecting behaviors against corporations and criminals, such as turning location sharing off or keeping accounts private, compared with their peers of color.

Discussion

Our survey asked young adolescents to report on their privacy behaviors, their endorsement of prescriptive privacy messages, and their privacy management preferences to understand how their desires for social connectivity and awareness of corporate surveillance and datafication practices shape their privacy management strategies. Results revealed that young adolescents' understandings and motivations related to digital privacy are complex and multifaceted, as their behaviors reflect trade-offs between privacy-protecting strategies against corporations or potential predators and disclosures to maintain social connectedness. On average, adolescents reported more privacy-protecting behaviors around social relationships than around corporations. They also endorsed prescriptive messages about predator-associated privacy risks more than they endorsed risks associated with corporate surveillance or future reputation.

Young teens were more likely to engage in privacy-protecting behaviors related to social relationships when they endorsed prescriptive predator messages and when they *did not prefer* risky peer relations over security. However, they were more likely to engage in privacy-protecting behaviors vis-a-vis corporations and criminals when they endorsed predator messages and *preferred* the convenience of corporate surveillance over security. Age was an important factor in our findings; 11- and 12-year-olds were less likely to practice privacy-protecting behaviors related to corporations and criminals but more likely to practice privacy-protecting behaviors related to social relationships compared to 13- and 14-year-olds. This finding aligns with other research observations—that when young adolescents begin using social media, their online social networks are small but expand as they age (Antheunis et al., 2016; Valkenburg et al., 2006).

The observed differences between the privacy beliefs, attitudes, and practices of young White adolescents and young adolescents of color may speak to manifestations of White privilege related to surveillance. Some research suggests that corporate surveillance is used in racialized ways, creating a digital manifestation of traditional surveillance tactics to perpetuate racism and anti-Blackness in the United States (Browne, 2015). Thus, it is perhaps unsurprising that White youth were less concerned than youth of color about sharing information with corporations and reported fewer privacy-protecting behaviors related to corporate data collection.

Our results highlight the importance of distinguishing between distal and proximal forms of privacy, as well as different forms of proximal privacy. Research finds that older adolescents and young adults try harder to present themselves favorably on social media (Dhir et al., 2016; Yau & Reich, 2019); thus, it is not surprising that the more proximal types of privacy related to social relationships loosen as adolescents get older, while their protective privacy behaviors against corporations and criminals may become more stringent. Young adolescents in our study also distinguished between different kinds of proximal privacy in peer relationships—they were less likely to endorse privacy preferences that introduced risk into their social relationships, yet they endorsed loose privacy preferences that facilitated peer social connectedness and relationships. These findings highlight how young adolescents' privacy preferences for different kinds of peer relationships cannot be easily grouped together, as only preferences for risky peer relationships predicted young middle schoolers' privacy behaviors. Communication privacy management theory (CPM) (Palen & Dourish, 2003) can help us understand how young adolescents learn that privacy is continuously negotiated and centered in peer interactions. The theory argues that tensions and "boundary turbulence" arise as people learn to navigate the complex boundaries of privacy, risk, and social connectedness/disclosure (Dourish & Anderson, 2006; Palen & Dourish, 2003). Vertesi et al. (2016), building on CPM theory, highlight the complicated intersection of boundaries youth navigate in the current digital age as they learn to balance their interpersonal relationships with people in their extended personal, professional, and consumer networks while "weighing a conflicting moral imperative to safeguard and protect their personal data and information disclosures" (p. 487). Our results directly speak to the tensions young adolescents experience, as they must weigh looser privacy preferences to facilitate greater peer connectivity, while not introducing too much risk to their social relationships (Wisnieski, Vitak, & Hartikainen, 2022).

Although young adolescents reported practicing a wide variety of privacy-protecting behaviors regarding corporations, potential predators, and social relationships, there were contradictions in their self-reported behaviors, beliefs, attitudes, and preferences. For example, adolescents who preferred corporate surveillance (i.e., agreed with statements such as, "I like seeing ads for things I like online") reported more privacy-protecting behaviors around corporation surveillance. Perhaps this is another kind of privacy paradox: that technology-savvy adolescents who know how to clear browser histories and block push notifications both appreciate the conveniences of, and are concerned about, corporate surveillance. Like the paradox described in the introduction, adolescents may be motivated by convenience, so much so that they are willing to give certain information to corporations for that benefit. We also found contradictions in young adolescents' beliefs and attitudes related to prescriptive messages for privacy protection from corporations. Youth largely agreed that corporations use your information to sell you things, but they tended to *disagree* that apps sell personal information to advertisers. These contradicting beliefs about prescriptive privacy messages could indicate that young adolescents do not understand the full scope of corporate surveillance tactics in their use of social media (Smith & Shade, 2018). Perhaps they view online spaces differently from smartphone apps. As young adolescents grow up in "networked publics," they must learn the details of corporate surveillance to inform their strategies for optimizing trade-offs for social connectedness and identity at a network level, where individual one-to-one strategies for managing interpersonal privacy become insufficient.

Marwick and boyd (2014) conceptualize a networked model of privacy, where achieving privacy for youth means developing "an understanding of and influence in shaping the context in which information is interpreted" (p. 1063). They emphasize the importance of controlling information flows on social media: that people can no longer entirely maintain one-to-one control over personal information. Instead, these choices and practices are networked, determined through a "combination of audience, technical mechanisms, and social norms," making privacy negotiation an ongoing process (p. 1062). Young adolescents are at a developmental moment when they are just beginning to learn to balance networked conceptualizations of privacy with interpersonal relationships in their daily lives. Importantly, their preference for some aspects of corporate surveillance, and their protective strategies against other datafication practices, may represent the turbulence or tensions that youth feel in weighing information disclosure and willing participation in corporate datafication practices with identity needs for autonomy, exploration, and social connectedness.

Social scientists have long theorized that the concept of privacy cuts across cultures, yet the ways in which privacy manifests itself culturally—how people *practice* privacy—is deeply contextual (Altman, 1977). Often, US legal models established to protect children and young adolescents (e.g., COPPA) conceptualize privacy in simplistic and individualistic models of human behavior (Cohen, 2012; Marwick & boyd, 2014), focusing on age 13 as an arbitrary cutoff for protections. However, privacy management in the current digital age occurs across vast networked publics and is highly culturally specific. Privacy management is both "contextual and relationally-accountable" (Vertesi et al., 2016, p. 479), and privacy is negotiated in social relationships that are embedded in culture; privacy management is neither a fully interpersonal nor fully networked process but a combination thereof. Therefore, understanding young adolescents' needs for social connectedness and identity exploration can help to explain the trade-offs adolescents make, namely disclosing personal information and becoming vulnerable to grow friend networks, gain connections, and express themselves. Rethinking how we implement age-gated protection measures for early adolescence, such as COPPA, will require a culturally situated approach that understands youths' social and identity needs within social media, instead of paternalistic protection measures that do not consider developmental needs for youth over age 13.

Additionally, understanding potential risks in digital spaces could be cognitively challenging to youth (Stoilova et al., 2019). Conceptualizing risk from others—such as a hacker who breaks into an account and steals private information or a sinister man pretending to be an eighth-grade female to gain one's trust—is easier than understanding that sites and applications are extracting data about social connections, geographic locations, and online activities. As such, youths' reasoning about risk and enacting behaviors that might mitigate those risks should be different from their reasoning for protecting against predators, criminals, and corporations.

Limitations

Because minimal work on digital privacy has focused on early adolescence as a developmental context, our research represents first steps in exploring adolescents' perspectives. Scale items were created to measure important distinctions between different types of privacy; however, some of these items did not hang together well as a single construct. For example, adolescents responded very differently to the two items meant to capture their privacy-protecting preferences for peer social relationships. Our data accounted for more variability in adolescents' privacy-protecting behaviors within social relationships compared to corporations and criminals; thus, future work should focus on better understanding young adolescents' perspectives and behaviors related to distal forms of privacy. In addition, our survey study was cross-sectional and therefore cannot identify causal relationships between attitudes, preferences, and behaviors. Future studies should use longitudinal designs to see how these factors relate over time. Additionally, it will be necessary to use experimental methods for studying young adolescents' privacy attitudes, preferences, and behaviors, as longitudinal data on their own are insufficient for identifying causal relationships between these factors. Further, we surveyed and compared only three grades (sixth, seventh, and eighth), so future work should explore how privacy practices change as students mature from grades six to eight. This is especially important as youth become more proficient social media users (Lenhart et al., 2011; Madden et al., 2013), as their social networks expand, as their need for identity formation and intimacy increase, and as federal protections like COPPA decrease. Another important limitation is that we asked what youth *do* to protect their privacy but did not *observe* what they do.

Our sample is composed of sixth through eighth graders at two large, East Coast United States middle schools where more than 50 percent of students identified as White and/or male. The demographics of our sample limit our ability to generalize these findings more broadly to other US samples and internationally. Though these findings shed some light on White students' preferences, they do not offer insights into the heterogeneity of practices for youth of color. Due to time constraints, some students were not able to complete the privacy preferences questions at the end of the survey. Therefore, our final sample likely represents students who were more conscientious, moved through the survey more quickly, or had a teacher who allowed more time for completing the survey during class. Notwithstanding these limitations, our study contributes to a limited body of work documenting young adolescents' digital privacy beliefs, preferences, and behaviors.

Conclusion

Children today are growing up in a highly connected world in which behaviors related to social relationships and consumerism are traceable. This constant surveillance by peers, corporations, and potential online attackers requires youth to consider their digital privacy at both a broader distal level and an interpersonal level. To date, most research has not included youths' voices and perspectives on privacy or considered how privacy attitudes and preferences relate to privacy-protecting practices. This study represents one step toward better understanding how the current generation of young adolescents conceptualizes digital privacy and how their privacy attitudes, beliefs, and endorsement of prescriptive privacy messages predict their privacy behaviors. Grounded in young adolescents' perspectives, our work casts a light on the perfect storm that adolescents are facing in the current digital age—weighing trade-offs between social connectedness, autonomy, identity exploration, and risky disclosure decisions.

References

Agosto, D. E., & Abbas, J. (2017). "Don't be dumb—that's the rule I try to live by": A closer look at older teens' online privacy and safety attitudes. *New Media & Society*, *19*(3), 347–365.

Albert, D., Chein, J., & Steinberg, L. (2013). The teenage brain. *Current Directions in Psychological Science 22*(2), 114–120.

Albert, D., & Steinberg, L. (2011). Judgment and decision making in adolescence. *Journal of Research on Adolescence*, *21*(1), 211–224.

Altman, I. (1977). Privacy regulation: Culturally universal or culturally specific?. *Journal of social issues*, *33*(3), 66-84.

Anderson, M., & Jiang, J. (2018). *Teens, social media & technology 2018*. Pew Internet and American Life Project. https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/

Antheunis, M. L., Schouten, A. P., & Krahmer, E. (2016). The role of social networking sites in early adolescents' social lives. *Journal of Early Adolescence*, *36*(3), 348–371.

Baumann, M. (2010). Pew report: digital natives get personal. *Information Today*, *27*(10), 18.

Blank, G., Bolsover, G., & Dubois, E. (2014). A new privacy paradox. In *Proceedings of the Annual Meeting of the American Sociological Association 2014*, 1–34.

boyd, d. (2014). *It's complicated: The social lives of networked teens*. Yale University Press.

boyd, d., & Hargittai., E. (2010). Facebook privacy settings: Who cares? *First Monday*, *15*(8).

Browne, S. (2015). *Dark matters: On the surveillance of blackness*. Duke University Press.

Caverlee, J., & Webb, S. (2008). A large-scale study of MySpace: Observations and implications for online social networks. In *Proceedings of the Second International Conference on Weblogs and Social Media* (ICWSM), 36–44.

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, *14*(2), F1-F10.

Cohen, A. O., Breiner, K., Steinberg, L., Bonnie, R. J., Scott, E. S., Taylor-Thompson, K., Rudolph, M. D., Chein, J., Richeson, J. A., Heller, A. S., Silverman, M. R., Dellarco, D. V., Fair, D. A., Galván, A., & Casey, B. J. (2016). When is an adolescent an adult? Assessing cognitive control in emotional and nonemotional contexts. *Psychological Science*, *27*(4), 549–562.

Cohen, J. E. (2012). *Configuring the networked self: Law, code, and the play of everyday practice*. Yale University Press.

Costello, C. R., McNiel, D. E., & Binder, R. L. (2016). Adolescents and social media: Privacy, brain development, and the law. *The Journal of the American Academy of Psychiatry and the Law*, *44*(3), 313–321.

Dhir, A., Pallesen, S., Torsheim, T., & Andreassen, C. S. (2016). Do age and gender differences exist in selfie-related behaviours? *Computers in Human Behavior*, *63*, 549–555.

Dourish, P., & Anderson, K. (2006). Collective information practice: Exploring privacy and security as social and cultural phenomena. *Human-Computer Interaction*, *21*(3), 319–342.

Federal Trade Commission. (1998). Children's Online Privacy Protection Rule (COPPA). https://www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/childrens-online-privacy-protection-rule

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience*, *26*(25), 6885–6892.

Gasser, U., & Palfrey, J. (2008). *Born digital: Understanding the first generation of digital natives*. Perseus Books Group.

Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *The British Journal of Sociology*, *51*(4), 605–622.

Harwell, D. (2020, April 1). Mass school closures in the wake of the coronavirus are driving a new wave of student surveillance. *The Washington Post*. https://www.washingtonpost.com/technology/2020/04/01/online-proctoring-college-exams-coronavirus/

Holloway, D., & Green, L. (2016). The internet of toys. *Communication Research and Practice*, *2*(4), 506–519.

Juvonen, J. (2007). Reforming middle schools: Focus on continuity, social connectedness, and engagement. *Educational Psychologist*, *42*(4), 197–208.

Juvonen, J., & Murdock, T. B. (1995). Grade-level differences in the social value of effort: Implications for self-presentation tactics of early adolescents. *Child Development*, *66*(6), 1694–1705.

Knifsend, C. A., & Juvonen, J. (2013). The role of social identity complexity in intergroup attitudes among young adolescents. *Social Development*, *22*(3), 623–640.

Lapenta, G. H., & Jørgensen, R. F. (2015). Youth, privacy and online media: Framing the right to privacy in public policy-making. *First Monday*, *20*(3). https://doi.org/10.5210/fm.v20i3.5568

Lenhart, A., & Madden, M. (2007). *Teens, privacy & online social networks: How teens manage their online identities and personal information in the age of MySpace*. Pew Internet and American Life Project.

Lenhart, A., Madden, M., Smith, A., Purcell, K., Zickuhr, K., & Rainie, L. (2011). *Teens, kindness and cruelty on social network sites: How American teens navigate the new world of "digital citizenship."* Pew Internet and American Life Project.

Livingstone, S. (2005). Mediating the public/private boundary at home: Children's use of the Internet for privacy and participation. *Journal of Media Practice*, *6*(1), 41–51.

Livingstone, S. (2006). Children's privacy online: Experimenting with boundaries within and beyond the family. In R. Kraut, M. Brynin, & S. Kiesler (Eds.), *Computers, phones, and the internet: Domesticating information technology*. *Human technology interaction series* (pp. 145–167). Oxford University Press.

Livingstone, S. (2008). Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Society*, *10*(3), 393–411.

Livingstone, S. (2014). Developing social media literacy: How children learn to interpret risky opportunities on social network sites. *Communications*, *39*(3), 283–303.

Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society*, *19*(5), 780–794.

Maalsen, S., & Dowling, R. (2020). Covid-19 and the accelerating smart home. *Big Data & Society*, *7*(2). https://doi.org/10.1177/2053951720938073.

Madden, M. (2017). *Privacy, security, and digital inequality*. Data & Society.

Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., Smith, A., & Beaton, M. (2013). Teens, social media, and privacy. *Pew Research Center*, *21*(1055), 2–86.

Marwick, A. E. (2008). To catch a predator? The MySpace moral panic. *First Monday*, *13*(6). https://doi.org/10.5210/fm.v13i6.2152

Marwick, A. E., & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, *16*(7), 1051–1067.

Marx, G., & Steeves, V. (2010). From the beginning: Children as subjects and agents of surveillance. *Surveillance & Society*, *7*(3/4), 192–230.

Miltgen, C. L., & Peyrat-Guillard, D. (2014). Cultural and generational influences on privacy concerns: A qualitative study in seven European countries. *European Journal of Information Systems*, *23*(2), 103–125.

Montgomery, K. C., Chester, J., & Milosevic, T. (2017). Children's privacy in the big data era: Research opportunities. *Pediatrics*, *140*(Suppl. 2), S117-S121.

Moscardelli, D. M., & Liston-Heyes, C. (2004). Teens surfing the net: How do they learn to protect their privacy? *Journal of Business & Economics Research (JBER)*, *2*(9), 43–56.

Palen, L., & Dourish, P. (2003). Unpacking "privacy" for a networked world. In R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, & G. Olson (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 129–136). ACM Digital Library.

Peter, J., & Valkenburg, P. M. (2011). Adolescents' online privacy: Toward a developmental perspective. In S. Trepte & L. Reinecke (Eds.), *Privacy online* (pp. 221–234). Springer. https://doi.org/10.1007/978-3-642-21521-6_16

Regan, P. M., & Steeves, V. (2019). Education, privacy, and big data algorithms: Taking the persons out of personalized learning. *First Monday*, *24*(11). https://doi.org/10.5210/fm.v24i11.10094

Rideout, V. J., & Robb, M. B. (2019). *The common sense census: Media use by tweens and teens*. Common Sense Media.

Robinson, S. C. (2016). iDisclose: Applications of privacy management theory to children, adolescents and emerging adults. In M. Walrave, K. Ponnet, E. Vanderhoven, J. Haers, & B. Segaert (Eds.), *Youth 2.0: Social Media and Adolescence* (pp. 139–157). Springer. https://doi.org/10.1007/978-3-319-27893-3_8

Rooney, T. (2012). Childhood spaces in a changing world: Exploring the intersection between children and new surveillance technologies. *Global Studies of Childhood*, *2*(4), 331–342.

Rudolph, M. D., Miranda-Domínguez, O., Cohen, A. O., Breiner, K., Steinberg, L., Bonnie, R. J., Scott, E. S., Taylor-Thompson, K., Chein, J., Fettich, K. C., Richeson, J. A., Dellarco, D. V., Galván, A., Casey, B. J., & Fair, D. A. (2017). At risk of being risky: The relationship between "brain age" under emotional states and risk preference. *Developmental Cognitive Neuroscience*, *24*, 93–106.

Shelley, M., Thrane, L., Shulman, S., Lang, E., Beisser, S., Larson, T., & Mutiti, J. (2004). Digital citizenship: Parameters of the digital divide. *Social Science Computer Review*, *22*(2), 256–269.

Shin, W., Huh, J., & Faber, R. J. (2012). Tweens' online privacy risks and the role of parental mediation. *Journal of Broadcasting & Electronic Media*, *56*(4), 632–649.

Shin, W., & Kang, H. (2016). Adolescents' privacy concerns and information disclosure online: The role of parents and the Internet. *Computers in Human Behavior*, *54*, 114–123.

Smith, A. R., Chein, J., & Steinberg, L. (2014). Peers increase adolescent risk taking even when the probabilities of negative outcomes are known. *Developmental Psychology*, *50*(5), 1564.

Smith, K. L., & Shade, L. R. (2018). Children's digital playgrounds as data assemblages: Problematics of privacy, personalization, and promotional culture. *Big Data & Society*, *5*(2). https://doi.org/10.1177/2053951718805214

Somerville, L. H., Hare, T., & Casey, B. J. (2011). Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *Journal of Cognitive Neuroscience*, *23*(9), 2123–2134.

Steeves, V., & Regan, P. (2014). Young people online and the social value of privacy. *Journal of Information, Communication and Ethics in Society*, *12*(4), 298–313.

Steeves, V., & Webster, C. (2008). Closing the barn door: The effect of parental supervision on Canadian children's online privacy. *Bulletin of Science, Technology and Society*, *28*(1), 4–19.

Steinberg, L., Cauffman, E., Woolard, J., Graham, S., & Banich, M. (2009). Are adolescents less mature than adults? Minors' access to abortion, the juvenile death penalty, and the alleged APA "flip-flop." *American Psychologist*, *64*(7), 583.

Stoilova, M., Nandagiri, R., & Livingstone, S. (2019). Children's understanding of personal data and privacy online—A systematic evidence mapping. *Information, Communication & Society*, *24*:4, 557–575. DOI: 10.1080/1369118X.2019.1657164.

Taylor, E., & Rooney, T. (2016). *Surveillance futures: Social and ethical implications of new technologies for children and young people*. Routledge.

Teräs, M., Suoranta, J., Teräs, H., & Curcher, M. (2020). Post-Covid-19 education and education technology "solutionism": A seller's market. *Postdigital Science and Education*, *2*(3), 863–878.

Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, *93*, 1–12.

Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology and Society*, *28*(1), 20–36.

Utz, S., & Krämer, N. C. (2009). The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *3*(2), Article 2. https://cyberpsychology.eu/article/view/4223

Valkenburg, P. M., Peter, J., & Schouten, A. P. (2006). Friend networking sites and their relationship to adolescents' well-being and social self-esteem. *CyberPsychology & Behavior*, *9*(5), 584–590.

Vertesi, J., Kaye, J., Jarosewski, S. N., Khovanskaya, V. D., & Song, J. (2016). Data narratives: Uncovering tensions in personal data management. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 478–490. https://doi.org/10.1145/2818048.2820017

Wisniewski, P. (2018). The privacy paradox of adolescent online safety: A matter of risk prevention or risk resilience? *IEEE Security & Privacy*, *16*(2), 86–90.

Wisniewski, P., Vitak, J., & Hartikainen, H. (2022). Privacy in adolescence. In B. P. Knijnenburg, X. Page, P. Wisniewski, H. Richter Lipford, N. Proferes, & J. Romano (Eds.), *Modern socio-technical perspectives on privacy* (pp. 315–336). Springer Open Access.

Yau, J. C., & Reich, S. M. (2019). "It's just a lot of work": Adolescents' self-presentation norms and practices on Facebook and Instagram. *Journal of Research on Adolescence*, *29*(1), 196–209.

Youn, S. (2009). Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents. *Journal of Consumer Affairs*, *43*(3), 389–418.

Youn, S., & Hall, K. (2008). Gender and online privacy among teens: Risk perception, privacy concerns, and protection behaviors. *CyberPsychology & Behavior*, *11*(6), 763–765.

Zuboff, Shoshana. 2019. *The age of surveillance capitalism*. Profile Books.

9

Humanizing Big Data

Making Sense of How Youth of Color Experience Personalized Educational Technologies

Veena Vasudevan

> "Why do we got to do this?" Zaire looks up at me with pained eyes as I walk over to see what he is working on that day. He has opened up Edgenuity to work on a humanities assignment about myths. "I don't want to read about myths!" I recognize the stressful relationship he has with complex text and the isolation of sitting in front of a computer to learn something that doesn't seem salient. We work slowly, reading, discussing the content, and eventually it starts to click. (November 6, 2015)

Students like Zaire are often on the receiving end of poorly conceived technological solutions touted to improve their learning and close achievement gaps. To explore how students engage with personalized educational technologies, I draw from a two-year ethnographic research study at an urban public high school, the Design School. Personalized or student-centered learning solutions are educational technologies that are being increasingly adopted by public schools and districts within the broader context of corporate school reform efforts (Roberts-Mahoney et al., 2016). Learning solutions that emphasize personalization are often misaligned with broader curriculum efforts in schools (Bingham et al., 2018) and emphasize the role of trackable data—referred to as data-driven decision-making—in shaping students' learning experiences (Roberts-Mahoney et al., 2016). They are rarely designed in conversation with students, educators, or parents, and they are often misaligned with the needs of youth from marginalized communities (Reich et al., 2017). Students' experiences with technology for the purposes of learning are increasingly mediated by educational technology platforms (or digital education platforms), which are digital applications with multiple functions (e.g., student data tracking, learning apps, communication tools) but provide one seamless experience for users and can integrate with other digital tools and applications (see Decuypere et al., 2021). Common platforms visible in schools include Canvas, Nearpod, and Peardeck.

These educational technology platforms apply logics to interpret personal preferences and make recommendations: in classrooms youth might engage with software that suggests books based on students' reading levels, previous selections, or assessment data. Politically, personalization allows districts to communicate that they value youth as learners and individuals within learning contexts. Practically, a shift toward "personalized" learning gives resource-starved districts a way to lessen costs and deal with increasing class sizes by outsourcing teaching and learning to online programs and prepackaged curricula (Basham et al., 2016; Bulger, 2016; Staker & Horn, 2012). However, personalization engines do not consider students' histories in person (Holland & Lave, 2001) or factors beyond the screen—whether a student had a bad bus ride into school, how to pivot when a student wants to embark on something new and adventurous, or how to give students opportunities to engage in dialogic learning experiences. Ultimately, these conceptions of "student-centered" learning draw on a framework of content transfer and delivery, or a dressed-up version of Freire's banking model of education (Freire, 1993). Personalized learning solutions are rapidly becoming a new kind of standardization: laptops instead of books, online quizzes in lieu of paper quizzes. They do not represent an inherent shift in *how* students learn so much as a shift in the medium by which students access content. Educational technologies like these espouse an instructionist approach to learning where technology is used for content delivery—learning *from* technology—versus a more constructionist approach, which encourages youth to learn through creating *with* technology (Kafai, 2006; Papert, 1993).

Drawing from the Design School students' experiences, I provide a foil for a larger argument about how personalized learning can reify gross inequities that already persist in public education. I pose the question, How do educational technologies shape the learning lives of students? Through examining one personalized learning technology, I hope to illuminate the practical and everyday ways that youth of color from marginalized communities experience technology, because it is important to humanize what big data can often obscure.

Conceptual Framing

In the United States, education has long reinforced social and economic hierarchies, limiting the quality and breadth of academic learning experiences in schools for youth of color.[1] Specifically, Black and Latinx youth have been relegated to second-class citizenship, tracked into technical and task-oriented careers that would benefit large-scale employers but not allow for the social mobility and change that schools promise (Anderson, 1988; Sanchez, 1993). Moreover, the educational system has been historically structured to silence, oppress, and demand compliance from youth in urban public schools (Anyon, 1981; Ferguson, 2020; Fine, 1991), resulting in educational environments that are often uncaring (Valenzuela, 1999; Vasudevan et al., 2022) and punitive (Ferguson, 2020) and that disproportionately discipline students rather than encouraging their academic achievement (Thomas & Stevenson, 2009). Moreover, schools in the United States have largely been used to reinforce differences in service of the political economy (Holt, 2002; Omi & Winant, 2014). Currently, particularly amidst a pandemic, the gross inequalities within our educational systems are laid even more bare.

There are a variety of ways in which students of color experience vast educational inequities. The design of curriculum, pedagogical practices, and the very nature of how knowledge is perceived are factors that shape students' academic lives in their everyday school experiences. Nieto (1999) argues that learning for many students of color gets reduced to "reproduction of socially sanctioned knowledge" (p. 3): students are positioned as empty vessels that require filling because they don't have the currency required to navigate school success, which implicitly justifies the banking model of education (Freire, 1993). Others have documented that formal education fails to recognize students' funds of knowledge (Moll et al., 1992), out-of-school literacies (Hull & Schultz, 2001; Vaughan, 2020), and capacities for learning and leadership. Black and Latinx students are also more likely to have less experienced educators, less academic technology and materials, and inadequate school facilities (Darling-Hammond, 2010), factors that are exacerbated by a culture of low expectations and an emphasis on standardized testing over authentic learning (Gadsden et al., 1996; Nieto, 1999).

One of the persistent equity challenges is in how students use educational technology in school. Scholarship on the digital divide attends to infrastructure (e.g., broadband connectivity), access (owning computers and mobile devices, opportunities to use them in school, etc.), and, increasingly now, deciphering *how* young people engage with technology in educational environments, as well as the quality and diversity of students' learning experiences *with* technology. Students in well-resourced schools have more access to sophisticated technologies, educational tools, increased teacher expertise, and professional facilities, rendering their experiences different from those in underresourced schools (Dolan, 2016; Warschauer & Matuchniak, 2010; Warschauer & Tate, 2018). Well-resourced schools can offer better and more varied educational opportunities that are built around more than simply passing standardized exams. Instructional technology in underresourced schools focuses disproportionately on drill and practice, memorization, and preparation for standardized tests, positioning students as consumers of technology versus producers (Dolan, 2016). While access to and availability of technology has increased, the opportunity to use technology to facilitate creativity, critical thinking, and problem-solving still eludes students in the most marginalized schools and communities (Rowsell et al., 2017; Warschauer & Tate, 2018). In his examination of adolescents' technology use in three middle schools, Rafalow (2020) found that White students were generally encouraged and even praised for inviting their technology practices into the classroom, while Black and Latinx students were discouraged or worse, disciplined for doing the same. Thus, even with access to potentially engaging and rich instructional technologies, students are not able to leverage these technologies in equitable ways (Rafalow, 2020). The digital divide is rightfully seen as an equity issue that goes beyond physical resources and infrastructure to reveal the ways in which using instructional technologies also reifies systemic educational inequities, with youth of color disproportionately affected.

What we also know is that when given access and opportunities, young people, including youth of color, draw from their lived experiences to creatively compose, cultivate, and nurture deep technical skills; produce a range of sophisticated multimodal artifacts; and engage in participatory cultures (Ito et al., 2010; Ito et al., 2013; Ito et al., 2020; Kafai & Peppler, 2011; Moje, 2000; Pinkard et al., 2020; Vasudevan, 2022). Many educational learning technologies have been developed specifically for tinkering, play, exploration, and learning, conceived through a constructionist approach to learning that engages students through design and problem-solving (Kafai, 2006; Resnick & Rosenbaum, 2013; Vossoughi & Bevan, 2014). These applications of technology are educational in that they facilitate youths' creative endeavors and fuel more creativity, questioning, and, in many cases, educational participation. Alternatively, the instructionist approaches of much of the personalization technology simply attempt to impart standards-aligned content; they are less concerned with engaging students in making, creating, and practicing *with* technology. Technology use in classrooms, particularly in underresourced schools, still reflects a very staid vision of school as a site for content acquisition (Dolan, 2016; Rafalow, 2020; Warschauer & Tate, 2018). This technocratic conceptualization of education suggests that simple proximity to technology will deepen learning, when in reality the transformative potential lies not just in access to technology but also the opportunities to engage with it creatively and playfully (Warschauer & Tate, 2018).

Methods

Site and Context

The Design School, located in a large metropolis in the Northeast, opened in the fall of 2014 during the maker-movement zeitgeist sweeping the educational reform landscape, which sought to support students in cultivating identities as makers and participate in maker-oriented educational activities (Honey & Kanter, 2013). The inaugural class had 99 students at the start of the school year. The student body was 82.8 percent African American, 14.1 percent Latinx 1 percent White, 1 percent Asian, and 1 percent other. Students with disabilities made up 13.1% of the student population, and 100 percent of the students were economically disadvantaged. Due to attrition, there were 89 students at the end of the first year.

In its second year, the demographics mirrored the year prior, while the number of students increased to 172 as the school added a new group of freshmen.

The founding school principal, Mr. Gilmore, was passionate about innovating on facets of teaching and learning that he felt had not worked in his nine years as a history teacher in the same district. Specifically, he strived to implement (1) student-centered learning, in which the school would prepare students for the real world by encouraging independence; (2) asynchronous learning, in which students could move at their own pace; and (3) competency-based grading, where students' work would be rooted in real-world (authentic) or "wicked" problems that focused on evidence-based measures of understanding and were a conscious shift away from what he described often as "learning from worksheets." To facilitate students' independence and self-paced learning, the school adopted a 1:1 laptop program, leveraging fairly inexpensive Google Chromebook laptops for each student. Beyond that, he and his small new staff were committed to encouraging youth leadership, employing restorative justice over antiquated and overused disciplinary measures, and supporting students' social-emotional development and growth.

While rich in an energetic staff and vision, the Design School, like most public schools in the district, lacked substantial resources. The school was colocated with another new high school in the same "innovation" network. Each school inhabited one floor in a former elementary school building that had been shuttered years prior, creating an odd juxtaposition—tall high school students walking the hallways of a school originally built for children under 10. The Design School's classrooms were peppered with a hodgepodge of mismatched desks, tables, and chairs, which Mr. Gilmore had collected from other district high schools that had closed.

In the first year (2014), the curriculum was designed and curated by the school's five teachers and was available via Google

Classrooms. In addition to Chromebooks, two of the seven classrooms were outfitted with outdated iMacs that lacked IT infrastructure—no unique user accounts, logins, or directories to save independent work. If students created a movie file, they could save it but with no guarantee that it would be there the next day. There were also no additive technologies like microphones, headphones, or other tools that would have allowed students to engage in creative production in a fully immersive way. The iMac software was out of date, and there was no comprehensive policy to maintain the machines. Nor was there any formal training for students (especially in the first year) on how to use any of the technology. There were other restrictions—YouTube and other social media sites were blocked, although students eventually figured out how to get around the district firewalls—and cell phones and personal devices went through a range of policies from totally banned to eventually being accepted in classrooms. None of the classrooms had projector screens, so slides, films, and other texts were projected onto small wall-mounted whiteboards or, in some classrooms, the wall. Some rooms didn't have permanent projectors, so floating projectors were wheeled in and out of the principal's office on carts that had also been sourced from closed schools.

During the school's second year, the administration and staff (both new and existing) began developing more systems and structures to support their model of asynchronous, student-centered, competency-based learning. The pedagogical structure for teaching and learning in the school's second year was composed of daily action plans that students used to set mini-goals in service of larger curricular goals, engaging in self-driven content via Edgenuity or Google Classrooms; intermittent mini-lessons in small groups; and one-on-one check-ins with teachers that students could schedule. Within the core disciplines (humanities, science, mathematics), students worked toward larger performance tasks: for example, writing an original myth in humanities. The Design School teachers collaborated across grades to select units and content they felt would align with their curricular goals and the larger themes they wanted to center (e.g., identity, conflict).

Edgenuity

Edgenuity is a web-based platform that school districts have adopted for a range of uses, including credit recovery, personalization, and increasingly, remote instruction (Edgenuity, 2014; Farmer, 2016; Llewellyn, 2019). The platform offers a slate of standards-aligned courses (linked to either the Common Core State Standards or state-specific standardized assessments). Edgenuity courses leverage video lectures, readings, formative assessments like quizzes (multiple choice or open ended), and more summative assessments like exams.

Edgenuity utilizes mastery approaches to learning in which students who demonstrate skills and conceptual knowledge can move to the next big topic. As such, the platform is designed to introduce topics using videos, short readings, and other multimedia formats, then offers formative assessments like multiple-choice or open-ended quizzes within each lesson. These assessments use keyword grading (Edgenuity Help Center, n.d.; Smith, 2020), which requires students to use specific words to get more than 0 percent. There are many types of assignments and scenarios, but most of the formative assessments require at least one keyword to get a score of 100 percent. The other elements of content delivery are customizable, and teachers can track students' progress and evaluate whether they have met sufficient expectations to move on or need to revisit the content (Edgenuity, n.d.; Eddy & Ballenger, 2016).

Edgenuity's model is didactic and does not facilitate dialogue. For example, there is no peer-to-peer engagement or even immediate feedback. Instead, students individually engage with content, respond to prompts, and receive automated grades from the system for interim assessments. Teacher-produced grades are received for more culminating assignments, but Edgenuity allows teachers to leave more (if not all) of a course's grading to the system (Edgenuity Help Center, n.d.). Moreover, because of the mastery approach, students have to return to the content to receive credit on formative and summative assessments, but it is not explicit how (if at all) the platform encourages critical thinking or problem-solving.

The Design School adopted Edgenuity to facilitate the type of asynchronous and self-paced learning that was critical to their model. Students' assignments on Edgenuity for mathematics and reading were informed by their scores on computerized adaptive tests called the Measures of Academic Progress (or MAP) that they took at the beginning of the 2015 school year. The scores were then used to create an individualized learning plan (ILP) that would be reflected in Edgenuity's MyPath, a "supplemental program that offers data-driven differentiated instruction for math and reading" (Edgenuity, 2014). The Design School adopted MyPath to help students "catch up in places they were behind and get ahead in places where they were strong" (student-facing artifact, 2015).[2] In addition, the school adopted the Keystone Biology curriculum from Edgenuity to supplement classroom instruction and help students prepare for state assessments. The school's offline curriculum wraparound was intended to bolster the online technologies like Google Classrooms and Edgenuity and to give students more structure. As the following student stories suggest, this was a work in progress that needed more refinement and attention to ensure that students felt connected to their learning experiences and rooted in a set of big ideas.

Positionality and Context

I came to the Design School as a member of a design-based ethnographic project that was examining the shifting nature of students' literacy practices. Our team collected data from three interdisciplinary labs that were part of the school's attempt to reimagine traditional urban public education. Our team collaborated with teachers to design learning experiences that were competency based, allowed for asynchronous engagement, and examined authentic problems. In addition to supporting the ethnographic research team, I was also embarking on my dissertation data collection, an educational ethnography that sought to explain the lived experiences of students at a maker-oriented high school. I wanted to focus on how learning that was asynchronous, student-centered, and maker-oriented might shift how students engaged with school. As I spent time with students and explored the ethnographic context, my focus shifted to helping students develop three youth-led spaces—a film club, a dance team, and a youth empowerment group—while also continuing to center their learning and literacy. Working with students in these out-of-school spaces of their own making allowed me to get to know them and observe their creativity, leadership, and technological savvy. I was also able to visit school classrooms and work closely with students on their academic work.

In my research, I wrote (Clifford & Marcus, 1986) and pictured (Ruby, 2000) culture by collecting a range of multimodal ethnographic data between 2014 and 2016, including ethnographic field notes, photographs, short films, and audio recordings. I also conducted semistructured interviews with students and staff and collected student artifacts and school materials (e.g., announcements, student memos, assignments). The students in my study were 13–15 years old when the study began in their freshman year (2014). I had informed consent from 27 students in the freshman class, of whom 6 became focal students, which allowed me to tell richer, more focused stories about their lived experiences in relation to schooling and identity. In this chapter, I highlight seven students' experiences, of whom three were focal students in my larger study (see table 9.1).

I primarily analyzed specific questions from students' transcribed interviews[3] about perceptions on teaching, learning, and educational technology. I brought students' interviews into conversation

Table 9.1
Student demographic data and their out-of-school interests

Student name	Age (at start of study)	Ethnicity	Youth-led space	Out-of-school/ outside-of-academic interests and passions
Denise	14	Latinx	Youth empowerment Group leader, focal student*	Writing novels, graphic design, video design, music
Tighe	14	African American	Film club	Writing and performing music, football team
Ruby	14	African American	Dance team Captain, focal student*	Writing novels, singing, poetry club
Aliya	14	African American	Dance team	Dance, cheerleading, doing hair
Anya	14	Latinx, Dominican	Youth empowerment Group leader, focal student*	Dream Project, Dominican culture, writing
Charles	14	Cambodian	Youth empowerment club	Cambodian culture, cooking
Aria	14	Latinx	Youth empowerment club	Music, anime

Note: * indicates cases where focal students and I worked together to create youth-led spaces and I spent substantive time with them in a range of informal and formal schooling contexts. We were actively in dialogue about my research.

with field notes that highlighted students' participation with educational and learning technologies (hardware and software) across both years of the study, as well as memos I had written while I was in the field. Ultimately, I utilized integrative and cross-conceptual memos (Emerson et al., 2011) to understand how educational and academic lives were shaped by technologies the students were required to use in the name of learning and academic advancement. Here, I analyze students' experiences with Edgenuity, which was implemented in their second year, revealing both challenges and concerns as well as the affordances educational web applications can have in shaping youths' lives.

Findings

The Edgenuity platform posed several challenges. Several students did not feel like they were learning anything—the experience felt like "answer-getting" and work completion. Some students felt bored and demotivated by the monotony of such a system, and others were frustrated that their actual knowledge and understanding was not reflected by their online program of study. For other students, asynchronous and online learning was effective because it reinforced the independence they craved, but even then the system did not deliver adequate learning experiences.

Work Completion over Learning

Tighe, a student I worked with through the youth-led film club, said Edgenuity left him feeling totally disconnected from his vision or expectation for school. He was passionate about music, football, and fitness. He was very close to his family and maintained a handful of strong relationships, including some at the Design School. Tighe was a tall, quiet, Black boy who would often located himself on the edges of the classroom, appearing at times to be aloof. But as time went on, he revealed himself to be a student who derived joy from active and embodied learning experiences. Early on he was skeptical of the Design School's approach to learning: "This whole computer thing, working on computer[s] and laptops is new for me still; even though I've been here for two years, I still can't—I can't do it; I can't work with them." When we started unraveling what specifically was challenging about using technology, Tighe offered, "Coz it's like I'm cheating myself when maybe don't test my mind . . . I can just—they give us 60 minutes to do a whole 10-question test—in 60 minutes, and all I gotta do is go on Google and check, it is gonna be there [the answers], and that's not right. Anybody could pass like that." Tighe was referring to the multiple-choice responses on the quizzes or assessments peppered throughout Edgenuity's digital curriculum. He explained further, "Yeah, it's the fact that I'm cheating myself out of my own education." Tighe was resistant to learning with computers. He didn't see himself or claim the identity as a "computer person," and learning with tools like Edgenuity felt disingenuous to what he thought learning should and could be.

During the same conversation, Tighe explained how his digital humanities lab class was a place where using technology was rewarding, saying that his teacher, Mr. Caulfield, "was teaching me things; I never knew about a camera, or I never knew nothing about iMovie, I never knew how to work it, so it was just like—it was fascinating how you can do these things with a computer." In contrast to how he felt about the Edgenuity experience, he believed using technology to create new content—like his music—and explore and learn about his interests was rewarding. His inclination toward the creative and performative was evident from the first time I spotted his spiral notebook of raps and rhymes, and it would later be evident during our work together in film club and via the student-led school talent show. Tighe embraced opportunities to take action, to try things out, and to participate in embodied learning experiences.

Charles shared a similar sentiment to Tighe, explaining that "I think only thing is just to do this and do that Edgenuity, and then I don't feel like I'm learning anything. All I'm doing is just taking notes and taking fake quizzes." Charles reflected that he wished high school was like the movies he watched where "teachers tell us to do the work instead of just assigning it, like force us to do it." As we chatted, he said of the online learning experiences, "Just like—just teach us instead of just having the computer as teachers." In the early stages of learning to work with Edgenuity, Charles and many other students expressed frustration that the teachers and the school weren't teaching. Charles did not have any issues navigating online applications or using a computer to complete assignments, but he did lament what felt like a negative shift—from learning with and from people to learning, as he suggested, from the computer.

Charles was an attentive student. He had up and down days but generally kept plodding along because he felt a sense of responsibility to graduate and do well. His aspirations to become a chef and his love for classic Cambodian oldies did not factor into his everyday work at school, but they certainly peppered his technology use. A quick scan of his Instagram Live or feed revealed all the food he sampled and classic musicians he reminisced about in his free time. Charles's admission that he wished teachers would just teach is one that many students expressed in the first two years at the Design School. This student discontent revealed a tension between the principal's vision for cultivating independence and the students' relearning of what it meant to learn at the school. In this case, however, Edgenuity as a resource to facilitate the vision for independent, asynchronous, and task-driven learning did not deliver.

Charles, Tighe, and many other students felt like they were cheating because it was too easy to look up answers and complete formative assessments like quizzes. While the performance tasks that culminated at the end of units were not things students could Google (in theory), it's clear that the processes of learning were facilitated by a system that made students feel like they were simply going through the motions. The implicit design of technologies like Edgenuity suggests that knowledge acquisition or answer-getting is the ultimate result. Such an approach centers a banking model of education (Freire, 1993) with a different affective experience—online quizzes instead of a sheaf of worksheets that require completion. Tighe, Charles, and others completed their work but only for the sake of finishing, not because they were challenged, inspired, or interested in where the content journey was meant to take them. Moreover, the Edgenuity design feature that Tighe critiqued, keyword grading, directly contributed to their feeling that they were doing busywork, not real learning.

Where they and other students were able to shine and make the work their own was in the performance tasks that teachers designed: Tighe wrote about music, and Charles focused on the Cambodian revolution for a culminating essay in humanities, allowing him to explore his cultural history. However, Edgenuity's delivery of the standards-aligned content did not facilitate this creativity or customization, and it did not always feel real. Instead, it was up to teachers to design these student-centered authentic assessments for learning.

Style Eclipses Connection in Student-Centered Learning

Many students at the Design School embraced the asynchronous and student-centered pedagogical approach to learning. A decentralized approach to teaching and learning was effective for these students because it gave them a sense of personal responsibility.

Ruby started attending the Design School three weeks into the first school year. She had been on the waitlist, but as the school experienced attrition in the early weeks, the principal called her and recruited her from the waitlist by, as she recounted it, speaking to her "egotistical side" by conveying that the Design School was a place she could pursue any dream she had. Inspired by his vision, she came to visit and immediately felt connected to the space and to the school's novel pedagogical approach.

Ruby described her previous educational experiences as frustrating, especially when she felt that teachers were "hovering over" her. She appreciated that the Design School's model of asynchronous learning and leveraging of technology gave her "a lot of responsibility." However, she also offered, "The thing about Edgenuity is it's very dull." She went on to say that she wished it incorporated dialogue or real-time opportunities to connect with the content providers:

> I wish they [the people in the online lessons] were actually there, like you know there's just videos of them, and like they put them on there and like they are talking and explaining and stuff. But like what if they were like actually like just Skyping and we actually have our conversation, like we all have our individual teacher, which sounds crazy, like there's a lot of teachers. But it like if we were like to do that and be able to actually talk to them like, "Hey, I don't get this; could you probably explain this again?" And he be like—and like the dude or the girl would be like, "Yeah, I could explain this again, or do you want me to just say like—" you know, stuff like that, I think that would be like way better. But you know you can't always do what you want.

Ruby is describing the dialogue that happens when you can be with educators in real time. She wanted to be able to ask a question, pursue a line of inquiry, or just clarify things in the moment. In other words, what was missing were opportunities to engage in dialogic learning (Freire, 1993). When opportunities to engage in critical conversations that help students explicate their understanding and practice critical thinking are unavailable, even carefully curated content can be reduced to answer-getting.

Ruby was fiercely independent. She embraced the culture of the Design School that emphasized students' agency and appreciated being left to do her own thing. At times that model worked for her, but other times it allowed her, and many other students who were seen as leaders at the school, to stray far away. There were long periods where these students weren't turning in lots of work or even showing up in class. Even though the wraparound curriculum was essential in rooting students, they were still getting lost in an environment of click-through text and videos.

Exacerbation of Systemic Inequalities in STEM Learning

Math interrupted At the Design School, mathematics went through some substantive challenges. During the school's first year, the first math teacher quit three months into the school year, leaving 90 students with no math instructor. He was unhappy with the school's approach and felt that he was engaged more as a disciplinarian than an instructor. A second teacher, with no prior teaching experience, was hired in January 2015 and quit after just two weeks. Finally, the administrators found a suitable replacement in mid-March, Ms. Capshaw. She stayed on until the fall of the following school year, only to leave midyear due to personal injury and a desire to pursue a career in leadership. Finally, in the school's third year (fall 2016), the administration found a mathematics teacher who was embraced by students and was instructionally strong. However, until this point, especially for students who were part of the inaugural class, math instruction was completely disrupted. Students were forced to rely on other teachers to mediate the online instruction while the school scrambled to keep the position staffed.

Many students' anxieties stemmed from their disrupted math education, with equal frustration arising from learning math via online modules and not working with teachers to deepen their understanding. Aliya, a student who swung from being very engaged in class to losing focus and progress, was one of many students who lamented Edgenuity's challenges, explaining that she missed instruction and being part of a larger class discussion. When we specifically discussed mathematics, she expressed frustration:

> Oh, we don't have a math teacher, nobody have a math teacher in ninth grade or 10th grade. So when Mr. T left, it's just like everything just start goin' behind, we doin' this Keystone thing or Edgenuity, and I feel like it don't help nobody. Like Edgenuity is like, everybody don't do it and then is like—I don't know. I just feel like we need a real teacher here for people to do math.

Aliya's frustration is completely understandable—clicking through an application without any structure or guidance is not an optimal learning experience. Watching videos or watching others solve problems is not adequate math education. The study of mathematics requires dialogue with ideas, opportunities for critical thinking, decomposing problems, and specifically examining and engaging with concepts through practice. Moreover, Aliya was also generally frustrated with a system that made her feel isolated in a subject that was central to high school success.

Online learning can create other challenges when students' experiences do not line up with their expectations. As a young Dominican woman determined to dream big, Anya's decision to attend the Design School was driven by her hope that it would be a place of possibility. Inspired by the school's vision to embrace students' interests and passions, Anya started her freshman year with enthusiasm and openness, willing to try new things and taking steps to be integral to the school community. Early on, her diligence and commitment to her schoolwork and her willingness to participate in school activities caught the attention of the school principal and her teachers. By her sophomore year, she fully embraced the identity as one of the leading students in her class, which was reinforced when friends and acquaintances asked her for help and when teachers asked her to mentor younger or less experienced students. However, Anya's anxieties around the future ran high. In her sophomore year, I was on my way out of Anya's advisory class when I noticed she seemed sad. In an excerpt from a field note, her concerns about the future are intertwined with frustrations about the school's learning technologies:

> I walk over to Anya and ask if she's okay. She responds, dragging out her words, "Yeahhhh, why, miss?" I mention she seems a little off. She says, "I'm just stressed!" and explains she is thinking about college. "Miss, I have a 3.78 GPA—is that good?" I reassure her that it is. She insists, "No, miss, I want your EXPEEERRRT TECH opinion." I tell her that she should keep her grades up and that extracurriculars matter. As we chat, it occurs to me students have

> just received their progress reports and so anxieties are high. She turns to her advisory teacher and asks her, "Miss Oswald, what'd you get on your SAT?" Miss Oswald nonchalantly replies, "Well, uh, let's see, I got a 690 in math and 6-something on my verbal, so whatever that is." Then, Anya turns to me. "Miss, what did you get?" I tell her my score and then explain, "It's different for you guys; they are changing the expectations and the scores are different as well." "Ohhh, okay. Mi-iiisss—we have to take the PSATs NEXT year! NEXT YEAR! And I only answered FOUR questions on this math exam, FOUR! How am I going to learn all of that by next year? And in Miss Santini's class, we are going over problem-solving but we are supposed to be doing algebra. Instead, Edgenuity makes you start from integers, adding/subtracting decimals. Like I know most of that, and while I'm doing it I learn some new things bu—tt, we are supposed to be doing algebra!" (Vignette, October 2015)

Anya and many other students were nervous about so many unknowns that high school poses: Will I get into college? What is the process? How do I prepare? Will I be able to keep up with the mathematics required? Even though she felt comfortable with certain mathematical ideas and content, the predesigned curricula dictated that she review the concepts again. Misalignments between how personalized technologies interpret students' learning and understanding can cause anxieties and frustration. In this case, Anya's MAP placement tests had brought her back to pre-algebra instead of where she felt she ought to be, in algebra.

The limits of learning technologies in science teaching Examining students' feedback about their learning experiences in relation to science was eye-opening. It clearly illustrated that innovative technologies that are intended to alleviate substantive educational inequities often miss the mark. Students at the Design School had a consistent science educator, Ms. Oswald, for the years I was there. However, Ms. Oswald did not have a budget that could support robust lab experiences, so students did not have opportunities to practice with science technologies and materials. Students instead participated in virtual labs in Edgenuity, where they were positioned as observers, not direct participants. This created tensions that are illuminated below by Denise and Aria, who offered critical insight into what was missing from their science learning experiences.

Denise, a young Latinx woman, embraced challenges and was independent and determined. She loved Ms. Oswald, her advisory and science teacher, and loved science. Her out-of-school life was full of creative pursuits—designing book covers and writing novels on Wattpad, obsessing over online tutorials for makeup and hair, and editing tribute videos of her favorite band, 21 Pilots. She started a youth empowerment group, spearheaded the school's first bake sale, and worked closely with me in managing the logistics of several school events. Denise offered this about her science learning:

> I like what I'm learning, I just don't like DOING it . . . I don't know how to explain it—I see it, I like it, okay—but then how can I just sit there on that thing and just sit and just watch, and watch and watch? I like HANDS-ON! BAM! I like if it's worth doing it. Like doing something. Not just sitting there watching. Like the LAB—we didn't get to do the lab! We had to freakin' click a BUTTON on Edgenuity and that was us doing the lab—like are you KIDDING ME? You lose interest in it. So you want the hands-on: the fun stuff! Let me cut open something, dude! Let's DO this.

Denise's creative and agentive practices outside of her academics were in stark contrast to how she had to pursue science learning via an online platform: she was limited to watching someone else navigate the lab versus doing it herself.

Another student, Aria, who secretly harbored a passion to be a pediatrician, mentioned that "instead of doing everything like online, like I want more of that hands-on approach" and listed off experiments that conjured images of "real," or authentic, science.

Embedded in both Denise's and Aria's characterizations of hands-on learning is a cry for authentic, embodied, and active learning experiences, which are often unavailable to students in underresourced schools (Darling-Hammond, 2010). Denise's frustration that their labs were just a click of a button illustrates the limitations of personalized educational technologies, particularly in subjects like earth science and biology. It is through practicing or doing science that students can explore personally meaningful phenomena, interface with domain-specific vocabulary, and grasp conceptual knowledge (Furtak & Penuel, 2019). These experiences also underscore the ways educational inequities faced by nondominant youth persist even with available technologies—because students often engage as consumers and passive participants in their learning despite access to technological devices (Dolan, 2016).[4]

Discussion

The Design School students' experiences illuminate key issues in how adopting personalized educational technologies can both reproduce and remedy educational inequities.

Personalized educational technologies like Edgenuity can remedy educational inequities because they are self-contained platforms that deliver standards-aligned content and assessments, freeing teachers to spend more personal time with students. Moreover, using these platforms gives students more autonomy because the platforms are self-paced and evaluate students on mastering content (Basham et al., 2016). This can relax pressure on educators to ensure that every child in the class is understanding and engaging with the same sets of ideas simultaneously. Personalized technologies also create opportunities for students to take more ownership of their learning. As Ruby shared, while the content and delivery mechanisms were not always appealing, she derived value from being able to make her own decisions about how she spent her time.

However, platforms like Edgenuity can also reproduce longstanding educational inequities. First, these platforms tend to be reductive: they position students as consumers of knowledge and liken learning to content acquisition. Second, these platforms replace social interaction with computer interaction. Third, they use limited data inputs to align students to the curriculum.

At the Design School, students' experiences using Edgenuity were frustrating because the learning process was flattened to content mastery, and the system was easily manipulated. Keyword tracking was one feature that contributed to the students' sense that they were not learning anything. Even the much-maligned ditto worksheet requires some teacher feedback, whereas Edgenuity's assessment feature did not require any human interaction—it can be automatically graded based on which keywords a student enters (Edgenuity Help Center, n.d.). True learning cannot be reduced to simple answer-getting, but with platforms like Edgenuity, this is effectively the outcome. Students who are required to use these tools for their academic achievement are positioned as mere consumers and recipients of content, reifying persistent inequities in the education system that relegate youth of color to less dynamic and critically engaging educational experiences (Warschauer & Tate, 2018). These personalization technologies take ownership and autonomy of learning away from teachers and students (Huis & Nagenborg, 2019).

Personalized learning technologies also tend to rely on algorithms and automation to replace the work teachers traditionally do in classrooms (Basham et al., 2016). Opportunities for dialogue, for thinking through problems out loud together, and for asking questions in real time are all features of in-person learning that aren't easily replicated with personalized technologies. There is, as Bulger (2016) deftly argues, a palpable sense that personal educational technologies would offer better instructional support to students. Even with available technologies, the quality of students' experiences vary tremendously based on students' race, class, and socioeconomic status and the culture of technology and learning in their schools (Rafalow, 2020). Using personalized learning technologies in already underresourced schools threaten to increase this divide.

The third issue relates to the dubious nature of what *personal* really means. As mentioned, Edgenuity, like similar platforms, suggests that it is a personalized educational technology that customizes learning experiences to students' needs. At the Design School, students' MAP test led to a system-generated individualized learning plan, which then informed what curriculum choices the students had available to them. All the customizations were predicated on one data input, which limits how customizable or adaptive the learning can really be (Bulger, 2016). This is why Anya was frustrated with starting at pre-algebra again—she felt ready to move on to algebra, but the data inputs suggested she was not. Even with sophisticated algorithms that could customize content to learners' needs, if the quality of the inputs is limited, then the learning experiences will be bounded by what the system can produce. The structures of Edgenuity limited how effective it could be in nurturing students' learning. In real time, the best educators in classrooms are aware of where students are developmentally by cultivating personal connections, learning about students, and adhering to an ethic of care and compassion. Systems that don't have authentic and continuous feedback loops cannot be characterized as "personalized" because a specific, one-time input cannot "know" the whole person (Basham et al., 2016). Knowing young people means you know about their ride to school that day or that their parents were going through a divorce, and you understand how that might impact their performance on a test or a quiz that day.

The Design School did not have a culture of low expectations. Conversely, teachers and staff were advocates for students and held them to high expectations. The school wanted students to become independent thinkers and owners of their learning journeys. Adopting personal learning technologies was one way to foster independence and ensure their student-centered model could be realized. The Design School designed around the technology platform, offering pedagogical solutions like mini-lessons and individual conferences for students, but the technology failed to inspire students; it was also frustrating and at times demoralizing for educators. While many students navigated their way through the content, many others struggled, often skipping or ignoring assignments until the end of term, when they then scrambled to finish a host of incomplete work. This rendered the whole experience as work completion and not as carefully curated and personalized learning experiences tailored to students' needs.

Algorithms are increasingly a significant part of our everyday lives, but we seemingly have little opportunity to push back or question the ways in which they can have an outsized influence (Willson, 2017). By adopting automated and algorithm-driven educational solutions to facilitate learning, we implicitly suggest that technology solutions are superior and that education as a discipline and the nature of teaching and learning is simply something technology can solve (Roberts-Mahoney et al., 2016).

In the nation's best schools, we expect students to be engaged in experiential or embodied learning experiences—they are not the exception but rather the rule. These schools possess well-resourced science classrooms where students have real equipment and machinery for scientific exploration, design classes with sophisticated software for graphic art and peer collaboration, powerful computers and robotics equipment to support students' interests in computer programming and engineering, and more. Many urban schools, like the Design School, are underresourced and forced to compete for every dollar and opportunity, despite the heroic efforts of teachers and staff. Technocratic solutions like Edgenuity are designed for credit recovery and content "acquisition." These approaches to learning harken back to the duality of public education that has persisted for over a hundred years, where youth of color, who are often located in urban contexts, receive education that positions them as secondary, as consumers versus producers, preparing for jobs that do not demand their intellect but rather their complacency, obedience, and silence (Anderson, 1988; Anyon, 1981; Fine, 1991).

Conclusion

Since the summer of 2020, countless articles and reports have lamented the "learning loss" as schools shuttered their physical doors due to the COVID pandemic (Wall & Franko, 2020). Parents with children as young as preschool aged were trying to negotiate online learning, new tools and technology platforms, and virtual classrooms where children were held accountable for seat-time for up to eight hours a day. Op-eds and research abounded on the toll that a lack of standardized testing would have on children who would fall further behind—especially the poorest and most vulnerable (Kuhfeld & Tarasawa, 2020; Reilly, 2020; Sparks, 2020). There is a genuine possibility that policy-makers will see the pandemic as an opportunity to automate educational functions and cede control to technologies that do not engage students in creative learning experiences or center human relationality in the learning process. Even in an environment like the Design School in its early years, where many students felt safe and connected to their peers and educators, introducing technologies as a replacement for meaningful teaching and learning relationships created challenges. Personalized educational technologies and platforms obscure the humanity and care that is central to real learning. To know children—to understand who they are, what matters to them, and how different disciplinary content and skills might animate and elevate their aspirations—is a very different goal from that of the pervasive curriculum in urban public schools. Instead, in an otherwise wealthy nation, the most vulnerable youths' learning experiences remain—even in the age of algorithms and sophisticated technologies—limited by a lack of understanding for what teaching and learning can and should be.

Personalized learning platforms driven by algorithms are at the forefront of corporate educational initiatives like the Amazon schools and the now defunct WeWork schools, as well as a new wave of philanthropic initiatives like the Chan Zuckerberg Initiative. We need to push back. Solutions that were marginally effective in the world of business and commerce should not impact millions of children, whose districts and leadership accept funding and parameters not out of choice but out of necessity. We must instead think critically about how we can learn from the creative and agentive ways youth make, create, share, and produce with technologies, and think carefully about how we can shift schools away from emphasizing students' roles as passively consuming technology and knowledge to actively creating, making, and producing *with* technologies.

Notes

1. I use the term *youth of color* to refer to students who do not identify as White and specifically youth who identify as Black, Latinx, Asian, and Indigenous students. Youth of color disproportionately experience educational inequities, such as emphasis on testing, limited opportunities for creativity, and extreme discipline, among others. The research I draw on more heavily emphasizes Black and Latinx students' experiences, so occasionally I will specifically refer to students who identify this way.

2. This references a document that students were given about Edgenuity, created by the school principal.

3. For this paper, I used denaturalized transcription for students' interview data (Bucholtz, 2000).

4. In this text, I use the term *nondominant* to represent those who identify as people of color (e.g., Black, Latinx, Indigenous, Asian); people who experience poverty; people with disabilities; individuals who identify as lesbian, gay, bisexual, transgender, or queer (LGBTQ+); those from immigrant communities; resettled refugees; and people who are English learners. Moreover, children, aged 0–18 are considered "vulnerable populations" when it comes to research. Therefore, I have come to identify "youth from nondominant communities" as those individuals aged 18 and younger who identify with one or more of the aforementioned identities.

References

Anderson, J. (1988). *The education of Blacks in the South*. University of North Carolina Press.

Anyon, J. (1981). Social class and school knowledge. *Curriculum Inquiry*, *11*(1), 3–42.

Basham, J. D., Hall, T. E., Carter, Jr., R. A., & Stahl, W. M. (2016). An operationalized understanding of personalized learning. *Journal of Special Education Technology*, *31*(3), 126–136.

Bingham, A. J., Pane, J. F., Steiner, E. D., & Hamilton, L. S. (2018). Ahead of the curve: Implementation challenges in personalized learning school models. *Educational Policy*, *32*(3), 454–489.

Bucholtz, M. (2000). The politics of transcription. *Journal of pragmatics*, *32*(10), 1439–1465.

Bulger, M. (2016). Personalized learning: The conversations we're not having (Working Paper 07.22.2016). Data and Society Research Institute. https://datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf

Clifford, J., & Marcus, G. E. (Eds.). (1986). Writing culture: The poetics and politics of ethnography: A School of American Research advanced seminar. University of California Press.

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. Teachers College Press.

Decuypere, M., Grimaldi, E., & Landri, P. (2021). Introduction: Critical studies of digital education platforms. *Critical Studies in Education*, *62*(1), 1–16.

Dolan, J. E. (2016). Splicing the divide: A review of research on the evolving digital divide among K–12 students. *Journal of Research on Technology in Education*, *48*(1), 16–37.

Eddy, C., & Ballenger, J. (2016). The effectiveness of an online credit recovery program on improving the graduation rates of students at risk of school failure. *School Leadership Review*, *11*(7). https://scholarworks.sfasu.edu/slr/vol11/iss1/7

Edgenuity. (2014, May 13). *Edgenuity MyPath™ | Edgenuity* [Video]. YouTube. https://www.youtube.com/watch?v=zRNvsBanmbU

Edgenuity Help Center. (n.d.). *Auto Grading Method Comparison*. https://help.edgenuitycourseware.com/hc/en-us/articles/360042418854-Auto-Grading-Method-Comparison

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes*. University of Chicago Press.

Farmer, M. R. (2016). Steps towards personalized learning using online asynchronous technology: A study of 7th, 10th, and 12th graders at a small rural school in Massachusetts [Doctoral dissertation, Lesley University]. Educational Studies Dissertations. https://digitalcommons.lesley.edu/education_dissertations/3

Ferguson, A. A. (2020). *Bad boys: Public schools in the making of black masculinity*. University of Michigan Press.

Fine, M. (1991). *Framing dropouts: Notes on the politics of an urban high school*. SUNY Press.

Freire, P. (1993). *Pedagogy of the oppressed*. Continuum International Publishing Group.

Furtak, E. M., & Penuel, W. R. (2019). Coming to terms: Addressing the persistence of "hands-on" and other reform terminology in the era of science as practice. *Science Education*, *103*(1), 167–186.

Gadsden, V. L., Smith, R. R., & Jordan, W. J. (1996). The promise of desegregation tendering expectation and reality in achieving quality schooling. *Urban Education*, *31*(4), 381–402.

Holland, D., & Lave, J. (2001). *History in person: Enduring struggles, contentious practice, intimate identities*. SAR Press.

Holt, T. C. (2002). *The problem of race in the 21st century*. Harvard University Press.

Honey, M., & Kanter, D. (2013). *Design. Make. Play. Growing the next generation of STEM innovators*. Routledge.

Huis, I., & Nagenborg, M. (2019). It's getting personal: The ethical and educational implications of personalised learning technology. *Journal of Philosophy in Schools*, *6*(1), 44–62.

Hull, G., & Schultz, K. (2001). Literacy and learning out of school: A review of theory and research. *Review of Educational Research*, *71*(4), 575–611.

Ito, M., Arum, R., Conley, D., Gutiérrez, K., Kirshner, B., Livingstone, S., Michalchik, V., Penuel, W., Peppler, K., Pinkard, N., Salen Tekinbas, K., Schor, J., Sefton-Green, J., & Watkins, S. G. (2020). *The connected learning research network: Reflections on a decade of engaged scholarship*. [Research-based policy report]. Connected Learning Allianc**e.**

Ito, M., Baumer, S., Bittanti, M., boyd, d., Cody, R., Herr, B., Horst, H. A., Lange, P. G., Mahendran, D., Martinez, K., Pascoe, C. J., Perkel, D., Robinson, L., Sims, C., & Tripp, L. (2010). *Hanging out, messing around, and geeking out: Kids living and learning with new media*. The MIT Press.

Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., Schor, J., Sefton-Green, J., Watkins, S. G. (2013). *Connected learning: An agenda for research and design*. [Research-based policy report]. Digital Media and Learning Research Hub.

Kafai, Y. B. (2006). Playing and making games for learning: Instructionist and constructionist perspectives for game studies. *Games and Culture*, *1*(1), 36–40.

Kafai, Y. B., & Peppler, K. A. (2011). Youth, technology, and DIY: Developing participatory competencies in creative media production. *Review of Research in Education*, *35*(1), 89–119.

Kuhfeld, M. & Tarasawa, B. (2020). *The COVID-19 slide: What summer learning loss can tell us about the potential impact of school closures on student academic achievement*. [Department of Education policy brief]. Northwest Evaluation Association [NWEA].

Llewellyn, T. (2019). *A program evaluation of student and teacher perceptions of an online Edgenuity high school course program in an urban high school* (Paper 1563898748) [Doctoral dissertation, William & Mary]. W&M ScholarWorks. https://doi.org/10.25774/w4-gwcx-jp54

Moje, E. (2000). "To be part of the story": The literacy practices of gangsta adolescents. *Teachers College Record*, *102*(3), 651–690.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice*, *31*(2), 132–141.

Nieto, S. (1999). *The light in their eyes: Creating multicultural learning communities*. *Multicultural education series*. Teachers College Press.

Omi, M., & Winant, H. (2014). *Racial formation in the United States: From the 1960s to the 1990s*. Routledge Press.

Papert, S. (1993). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Peppler, K. A., & Kafai, Y. B. (2007). From SuperGoo to scratch: Exploring creative digital media production in informal learning. *Learning, Media and Technology*, *32*(2), 149–166.

Pinkard, N., Martin, C. K., & Erete, S. (2020). Equitable approaches: Opportunities for computational thinking with emphasis on creative production and connections to community. *Interactive Learning Environments*, *28*(3), 347–361.

Rafalow, M. H. (2020). *Digital divisions: How schools create inequality in the tech era*. University of Chicago Press.

Reich, J., Ito, M., & Team, M. S. (2017). *From good intentions to real outcomes*. Connected Learning Alliance. https://clalliance.org/wp-content/uploads/2017/10/GIROreport_v3_complete.pdf

Reilly, K. (2020, December 8). The learning gap is getting worse as schools rely on remote classes, especially for students of color. *Time*. https://time.com/5918769/coronavirus-schools-learning-loss/

Resnick, M., & Rosenbaum, E. (2013). Designing for tinkerability. In M. Honey & D. Kanter (Eds.), *Design, Make, Play* (pp. 163–181). Routledge.

Roberts-Mahoney, H., Means, A. J., & Garrison, M. J. (2016). Netflixing human capital development: Personalized learning technology and the corporatization of K-12 education. *Journal of Education Policy*, *31*(4), 405–420.

Rowsell, J., Morrell, E., & Alvermann, D. E. (2017). Confronting the digital divide: Debunking brave new world discourses. *The Reading Teacher*, *71*(2), 157–165.

Ruby, J. (2000). *Picturing culture: Explorations of film and anthropology*. University of Chicago Press.

Sanchez, G. J. (1993). *Becoming Mexican American: Ethnicity, culture, and identity in Chicano Los Angeles, 1900–1945*. Oxford University Press.

Smith, A. (2020, September 3). Students uncover method to manipulate grading algorithm. *The Independent*. https://www.independent.co.uk/life-style/gadgets-and-tech/news/students-ai-grading-algorithm-edgenuity-keyword-spam-a9703751.html

Sparks, S. (2020, April 9). Academically speaking, the "COVID Slide" could be a lot worse than you think. *Education Week*. https://www.edweek.org/leadership/academically-speaking-the-covid-slide-could-be-a-lot-worse-than-you-think/2020/04

Staker, H., & Horn, M. B. (2012). *Classifying K-12 blended learning*. Innosight Institute.

Thomas, D. E., & Stevenson, H. (2009). Gender risks and education: The particular classroom challenges for urban low-income African American boys. *Review of Research in Education*, *33*(1), 160–180.

Vasudevan, V. (2022). Designing their own curriculum: How youth co-constructed a dance team that opposed traditional student–school relationships. *Curriculum Inquiry*, *52*(1), 9–30.

Vasudevan, V., Gross, N., Nagarajan, P., & Clonan-Roy, K. (Eds.). (2022). *Care-based methodologies: Reimagining qualitative research with youth in US schools*. Bloomsbury Publishing.

Vaughan, A. (2020). Conceptualizing scholarship on adolescent out-of-school writing toward more equitable teaching and learning: A literature review. *Journal of Adolescent & Adult Literacy*, *63*(5), 529–537.

Vossoughi, S., & Bevan, B. (2014). Making and tinkering: A review of the literature. *National Research Council Committee on Out-of-School Time STEM*, *67*, 1–55.

Wall, P., & Franko, K. (2020, October 30). On pandemic "learning loss," schools look forward, not back. *Associated Press*. https://apnews.com/article/newark-schools-coronavirus-new-jersey-0dd87cdaeb60aeb9f2a5a011867f22a5

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education*, *34*(1), 179–225.

Warschauer, M. & Tate, T. (2018). Digital divides and social inclusion. In K. Mills, A. Stornaiuolo, A. Smith, & J. Zacher Pandya (Eds.), *Handbook of writing, literacies, and education in digital cultures* (pp. 63–75). Routledge.

Willson, M. (2017). Algorithms (and the) everyday. *Information, Communication & Society*, *20*(1), 137–150.

10
The 4 As: Ask, Adapt, Author, Analyze

AI Literacy Framework for Families

Stefania Druga, Jason Yip, Michael Preston, and Devin Dillon

Introduction

Children in the current digital information era are rapidly engaging with technologies powered by artificial intelligence (AI). AI refers to the intelligence possessed by machines, thus why it is also known as machine intelligence. Unlike humans, machines acquire intelligence through algorithmic techniques inspired from domains like statistics, mathematical optimization, and cognitive science, and they are fueled by computer processing power and a large amount of data (Legg & Hutter, 2007). AI systems show great promise in helping children and families improve online search quality, increase accessibility to internet search via advances in digital voice assistants, and promote AI-supported learning (Grossman et al., 2019; Ruan et al., 2020; Ruan et al., 2019). However, AI systems can also amplify bias, sexism, racism, and other forms of discrimination, particularly for those in marginalized communities (Angwin et al., 2016; Buolamwini & Gebru, 2018). Promoting critical understanding of AI—or AI literacy—for children and families is essential in this context.

Without AI literacy, families, mainly from historically marginalized groups, risk falling prey to misinformation and fear; they also risk missing potential opportunities for learning (Ferguson, 2012; Gebru, 2019; O'Neil, 2016). Families and children must work together to learn about AI systems and to think critically about how this technology impacts their lives (Druga et al., 2019). Prior research on family engagement with digital technologies stresses how important it is to consider variation among families and parenting styles (Coyne et al., 2017; Takeuchi & Stevens, 2011). Therefore, to support algorithmic justice in families, we need to consider how diverse families can access these skills (DiSalvo et al., 2016; Yardi & Bruckman, 2012).

AI literacy does not occur in a vacuum but is influenced by social, cultural, institutional, and techno-infrastructural contexts. We need to consider the ecological and situational issues surrounding families and how macrofactors and microfactors influence AI literacy in the modern family. Therefore, it is crucial to address the socio-ecological conditions that influence how families may adopt AI literacy and to create guidelines that integrate human-centered design into practice. An analysis of ecological systems (Bronfenbrenner, 1994) can explain how families could succeed with AI literacy; it can also unveil the broader implications of such an intervention. There is a parallel need to develop design practices and frameworks that support the development of systems encouraging equitable and informed understandings of the creation and use of AI (Gonzales, 2017).

Research on how families interact with home technologies is a growing area, providing implications for the design of new smart devices (Druga, 2018; McReynolds et al., 2017). Studies demonstrate that families can play a decisive role in guiding children on how to make meaningful use of technologies (Ito et al., 2009; Stevens & Penuel, 2010; Takeuchi & Stevens, 2011). However, the rapidly changing digital landscape is making it difficult for families to integrate advanced technology in meaningful and intentional ways.

Limited knowledge exists on how parents or guardians learn with their children using tools that promote AI literacy. We wish to advance this body of research by posing the following research questions:

How do children and parents from different countries and diverse socioeconomic statuses (SES) perceive and interact with AI?

How can we best support parents to scaffold their children's use of AI technologies in the home?

How can we design future technologies to best support families' AI literacy?

Our goal is to understand how to facilitate AI literacy in families better. We investigate this from two perspectives: an ecological evaluation of current AI systems and the design of new systems for AI literacy. Our research puts forth both a conceptual and empirical understanding of how families engage with AI literacy activities. Such an understanding can inform the design of culturally tailored tools and resources. We contribute new insights on family AI practices to address critical AI literacy needs in families. Finally, we develop a foundation that can encourage innovations to take advantage of family dynamics in a way that improves AI literacy learning. We analyze and compare different prior data sets to propose a novel, research-based, family-facing framework for thinking with and about AI.

We begin with a brief review of ecological systems that support AI literacy (Bronfenbrenner, 1994). Ecological systems theory refers to the nested systems—macrosystems, exosystems, mesosystems, and microsystems—that influence the development of learning for people in the following ways:

- Macrosystem factors: Social and cultural values
- Exosystem factors: Technology infrastructure and policies
- Mesosystem factors: Community centers, libraries, and schools
- Microsystem factors: Families, peers, siblings, extended family, and neighbors.

Through a review of the literature, we consider how current technological systems are supporting or not supporting the development of AI literacy. From our evaluation of ecological systems in AI literacy, we inductively develop a design framework that supports critical understanding and use of AI for families. Our framework considers four dimensions of AI literacy: *ask*, *adapt*, *author*, and *analyze*. We prototype and refine different learning activities such as detecting bias, testing a voice assistant, coding a smart game, and drawing what is inside the smart devices to explain how they work. These activities took place during four co-design sessions with an intergenerational group, consisting of adult design researchers, child participants ($n = 11$, ages 7–11 years old), and parents. The activities correspond to the different dimensions of our AI literacy framework.

Through a series of family co-design sessions, we found that children perceive bias in smart technologies differently from adults, and they care less about technological shortcomings and failures as long as they are having fun interacting with the devices. Family members supported each other in various collaborative sensemaking practices during the sessions by building on each other's questions, suggesting repairs for communication breakdowns with the voice assistants, coming up with new and creative ways to trick the AI devices, and explaining or demonstrating newly discovered features.

We demonstrate how our novel framework supports AI literacy development through play, balanced partnership, and joint family engagement with AI learning activities, concluding with a series of guidelines for families.

Finally, we engage in a broader discussion that connects the ecological systems theory with our AI literacy framework to draw implications for the broader perspective of AI practice, program design, public policy, and algorithmic justice.

The Ecology of Family AI Literacy

Based on our evaluation of ecological systems (Bronfenbrenner, 1994), we discuss the impact of multiple nested systems (i.e., macrosystems, exosystems, mesosystems, and microsystems) on family AI literacy.

Macrosystem Factors: Sociocultural Values

Fostering an environment where different identities can flourish Macrosystems impact learning and technology practices within values, policies, and infrastructure (Bronfenbrenner, 1994). One macrosystem factor in AI literacy is the importance of promoting an inclusive AI education for multicultural and multilingual families from different socioeconomic backgrounds. This approach requires us to consider diverse families outside WEIRD populations (i.e., Western, educated, industrialized, rich, and democratic; see Henrich et al., 2010). To include multiculturalism as a macrosystem factor for AI education, we need to be reflexive and consider how researchers approach such issues (Schön, 1987). We also recognize that, as Medin and Bang (2014) describe, the answers to our research questions will be influenced by the sociocultural values of the person "who is asking." We build on prior work on technology literacy and joint media engagement among multicultural families (Banerjee et al., 2018; Pina et al., 2018). As we conceptualize AI literacy, we define the term *literacy* as practicing rather than developing one's skills (Cole et al., 1997; Kulick and Stroud, 1993; Scribner and Cole, 1981). We situate the AI literacy practice in the constellation of sociocultural practices that our families engage in (Rogoff et al., 2014). In our effort to discover, encourage, and promote best practices of families using AI technologies in meaningful ways, we acknowledge the need to recognize multiple literacies and the relationships of power they entail (Street, 2003). Therefore, we seek to foster an environment where heterogeneity, specifically different identities, goals, and forms of learning and growth, can flourish (Rosebery et al., 2010).

Exosystem Factors: Technology Infrastructure and Policies

The brave new world of connected homes Necessary technological infrastructure also determines access to AI literacy. For instance, a 2019 Pew study shows that in the US, broadband access is limited by data caps and speed (Anderson, 2019). As AI systems increasingly take advantage of large-scale technological infrastructures, more families may be left disengaged if they cannot connect to broadband (Riddlesden & Singleton, 2014). Moreover, it is essential for minority groups to not only "read" AI but also to "write" AI. Smart technologies do much of their computing in the cloud, and without access to high-speed broadband, marginalized families will have difficulty understanding and accessing AI systems (Barocas & Selbst, 2016). Families must be able to use AI systems in their homes so they can develop a deeper understanding of AI. When designing AI education tools and resources, designers need to consider how the lack of access to stable broadband might lead to an AI literacy divide (van Dijk, 2006).

Policies and privacy Risks to privacy are standard on the internet. Studies show that privacy concerns constitute one of the main worries among children in Europe (Livingstone, 2018; Livingstone et al., 2011; Livingstone et al., 2019), and adults widely support the introduction of data protection measures for youth, such as Article 8 from the EU's General Data Protection Regulations (GDPR) (Lievens, 2017; Regulation (EU) 2016/679 of the European Parliament and Council, 2016). According to a recent survey, 95 percent of European citizens believe that "under-age children should be specially protected from the collection and disclosure of personal data," and 96 percent think that "minors should be warned of the consequences of collecting and disclosing personal data" (European Commission, 2011).

Furthermore, many companies do not provide clear information about the data privacy of voice assistants. In this context, policymakers and technology designers must consider the unique needs and challenges of vulnerable populations. Normative and privileged lenses can impair conceptualizations of families' privacy needs while reinforcing or exacerbating power structures. In this context, it is crucial to provide updated policies that look at how the AI technologies embedded in homes not only respect children's and families' privacy but also account for future potential challenges.

For example, the Children's Online Privacy Protection Act (COPPA), which passed in the US in 1998, seeks to protect kids under the age of 13. Despite the proliferation of voice computing since then, the Federal Trade Commission did not update its COPPA guidance for businesses until June 2017 to account for internet-connected devices and toys. COPPA guidelines now state that online services include "voice-over-internet protocol services" and that businesses must get permission to store a child's voice (Federal Trade Commission, 2017). However, recent investigations have found that in the case of the most widely used voice assistant, Amazon's Alexa, only about 15 percent of "kid skills" provide a link to a privacy policy. Particularly concerning is the lack of parental understanding of AI-related policies and their relation to privacy (McReynolds et al., 2017). While companies like Amazon claim they do not knowingly collect personal information from children under 13 without the parent's or guardian's consent, recent investigations prove that is not always the case (Lau et al., 2018; Zeng et al., 2017).

Nonprofit organizations such as Mozilla, Consumers International, and the Internet Society have since decided to take a more proactive approach to these gaps by creating a series of guidelines that teach families how to better protect their privacy (Rogers, 2019). These efforts could be used to increase AI literacy by helping families understand what data their devices are collecting, how these data are being used or potentially commercialized, and how they can control their devices' privacy settings or require access to such controls when they do not exist.

Mesosystem Factors: Community

Mesosystem factors refer to interactions in one setting that can influence the interactions in another setting. For instance, what happens in a library, school, or community center for children and families can influence learning at home (and vice versa). Studies show parental involvement in learning at home significantly influences school performance (Barron, 2004; Berthelsen & Walker, 2008) and can be critical to children's future success. For instance, the AI Family Challenge (AIFC) was a 15-week program implemented with third- through eighth-grade students ($n = 7,500$) and their families in underresourced communities across 13 countries. During the program, families learned to develop AI-based prototypes that solved problems in their communities. The goal of AIFC was to determine whether AI was of interest to such communities and to determine the impact of such intervention on participants' AI literacy. To gain insight into these objectives, researchers conducted pre-program and post-program surveys as well as interviews with participants in the US, Bolivia, and Cameroon (Chklovski et al., 2019).

After AIFC, 92 percent of parents believed their children could better explain AI to others, and 89 percent believed their children were capable of creating an AI application. The study findings indicated the need to improve parent training materials, connect technical mentors to local sites, and improve the curriculum to be more hands-on, engaging, and better illustrative of machine learning concepts.

Microsystem Factors: Families, Peers, Siblings, Extended Family, and Neighbors

Microsystem factors refer to specific interactions within the local environment that influence family learning. For this review, we look closely at family interactions in the home around AI literacy.

An example of these sorts of family interactions, from Technovation, can be seen in figure 10.1.

A survey of 1,500 parents of elementary and middle school students, commissioned by Iridescent Technovation (2019), found that 80 percent of parents in the US believe AI will replace most jobs (not just low-skilled jobs), less than 20 percent understand where and how AI technologies are currently used, 60 percent of low-income parents have no interest in learning about AI, and less than 25 percent of children from low-income families have access to technology programs (Chklovski et al., 2019). Research on families' interactions with technology is a growing area, providing implications for the design of new agents (McReynolds et al., 2017). As devices become more humanlike in form or function, humans tend to attribute more social and moral characteristics to them (Druga, 2018; Druga et al., 2018; Kahn et al., 2011; Kahn, Jr., et al., 2012). These findings raise the question of how parents need to engage and intervene in children's interactions with connected toys and intelligent agents. Studies show that parents scaffold their

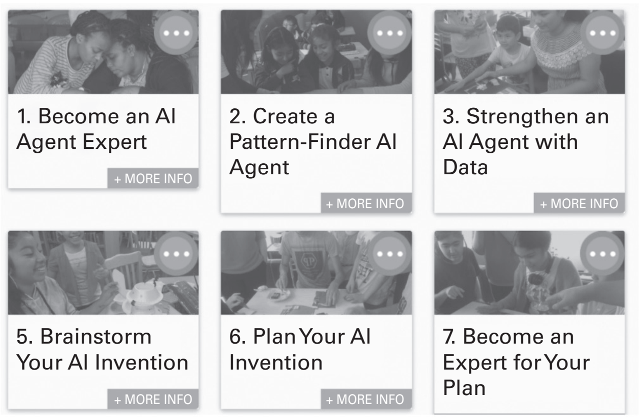


Figure 10.1
Example of curriculum modules created by Technovation for the international Curiosity Machine Competition for families.

children's behavior when the family interacts with robots or interactive devices together (Lee et al., 2006). We observed the same behavior when families interact with voice user interfaces (VUIs), particularly when parents help children repair various communication breakdowns with the conversational agents (Beneteau et al., 2019; Druga et al., 2017; Lovato & Piper, 2015). For instance, Beneteau and her colleagues (2019) noted that family interactions with Amazon Alexa devices facilitated joint media engagement conversations with parents. At the same time, however, the devices could not "code switch" between adult and child requests. This led to many frustrations and ultimately communication breakdown between the families and the voice assistant. In a longitudinal study analyzing families' uses of VUIs in the home, Porcheron et al. also showed that collaborative information retrieval is prevalent (2018). Both children and parents use classical conversation techniques, such as prosody changing or strategic use of silences, even if they dialogue with a more transactional agent like Alexa (Beneteau et al., 2019).

Methodology

Through our analysis of the ecological perspective on the state of AI understanding for families, and building on theories of parental mediation and joint media engagement (Takeuchi & Stevens, 2011), we propose a new framework for defining family AI literacy. To examine our framework in action, we adhere to the standards and practices of participatory design (PD), precisely the method of cooperative inquiry (Druin, 2000; Guha et al., 2004). Under Cooperative Inquiry in PD, adults and children work closely together as design partners, emphasizing relationship building, cofacilitation, design-by-doing together, and idea generation (Yip et al., 2017). Cooperative Inquiry works well for understanding AI systems and literacy because children already work closely with adults and are more likely to express their perceptions around childhood (Woodward et al., 2018). In design partnerships, there is a strong emphasis on relationship building, which allows children to be more receptive to experimentation and open dialogue.

Our co-design sessions focused on designing and eliciting responses from children and families around their perceptions of different aspects of AI systems. We conducted three 90-minute sessions from October to November 2019 with eight to 11 children. We also worked with families in co-design sessions in December 2019 to understand children's engagements with AI with their parents.

Participants

An intergenerational co-design group, consisting of adult design researchers (undergraduates, master's, and doctoral students) and children ($n = 11$, ages 7–11), participated in the four design sessions. The team was called KidsTeam UW (all names within the team are represented as initials). Children typically participated in the study for 1–4 years (2016–2019). In the fourth session, three KidsTeam UW children and their families (e.g., parents, siblings) came on a weekend co-design session to engage together and discuss their perceptions of AI technologies.

Design Sessions

Each KidsTeam UW design session (both child and families) consisted of snack time (15 minutes), where the children gathered to eat, share, and develop relationships through play. In circle time (15 minutes), we provided children a "question of the day" to prime them to think about the design session. We also provided the instructions (verbally and through activity printouts) for engagement. Most time was spent designing together (45 minutes), in which children participated in some design techniques (Walsh et al., 2010; Walsh et al., 2013; Walsh & Wronsky, 2019) with at least one adult partner. Children then broke up into smaller teams or remained together for a single design activity. Finally, the group came back together in discussion time (15 minutes) to reflect on the design experience.

We organized the sessions in the following way to investigate how the family AI literacy framework could be utilized as a series of design activities:

Design Session 1 (October 2019): We showed the children different video clips of "algorithmic bias." Video clips included AI not being able to recognize darker skin tones, voice assistants stuck in an infinite loop, and a very young child unable to get an Alexa Echo device to start. We used big paper (Walsh et al., 2013), a technique that allows children to draw on large sheets of paper to consider what "bias" means.

Design Session 2 (October 2019): We provided children with different technology activities using three kinds of AI devices: Anki Cozmo (AI toy robot), Alexa Echo voice assistant, and Google Quick, Draw! (AI that recognizes sketches). Each intergenerational team went through the stations and documented what was "surprising" about the technology and whether they were able to "trick" the AI system into doing something unexpected.

Design Session 3 (November 2019): Using big paper, we asked children and adults to draw how they thought a voice assistant (Amazon Alexa) worked.

Design Session 4 (December 2019): Finally, five KidsTeam UW families came together on a weekend morning workshop to engage in multiple AI technologies stations. Stations included Amazon Alexa, Google Quick, Draw!, and the Teachable Machine. One station used Cognimates (Druga, 2018) and BlockStudio (Banerjee et al., 2018) to show models on how computers made decisions. Families spent, on average, 15 minutes per activity trying out the different technologies and then wrote their ideas and reflections on the technologies.

Data Analysis

We used an inductive process to analyze the themes captured from the audio of family AI interactions (Charmaz, 2006). We began with memoing and open coding during the initial transcriptions of the video files. Through memoing and open coding, we noticed emerging themes related to family AI literacy practices and family joint engagement. We then began coding literacy practices and joint engagement from transcripts of each of the five families, developing and revising codes as we found additional examples of AI joint engagement, reviewing a total of 17 hours of video capture. We continued this process until codes were stable (no new codes were identified) and applicable to multiple families. Once the codes were stable, we again reviewed transcripts from each of the five families for AI literacy practices and family joint engagement. We included AI literacy practices from each participant in our corpus of 350 AI family–AI interactions, systematically going through each family's transcript and pulling out for each code (when present). For our final analysis of each family's AI interaction, a total of 180 AI interactions falling under the broad themes of AI literacy practices were deeply analyzed by two researchers. We defined AI literacy practices as interactions between family members and the various AI technologies, as shown in table 10.1. We drew on the human-computer interaction conversational analysis approach to analyze family interactions in an informal learning environment, with a focus on the participants' experiences.

AI Literacy Dimensions: The 4 As

Based on our analysis of the ecological perspective (Bronfenbrenner, 1994) of the state of AI and building on our prior work (Beneteau et al., 2019; Druga, 2018; Druga et al., 2017; Druga et al., 2018;

Druga et al., 2019), we consider ways to connect design dimensions for family AI literacy. Building on parental mediation and joint media engagement frameworks (Takeuchi & Stevens, 2011), we aim to analyze and support the scaffolding parents might provide to enable their children's mental models of intelligent systems. In this section, we highlight our novel framework for family AI literacy (see table 10.1) based on a thorough examination of the literature and our inductive co-design study. Our framework is composed of four dimensions (4As)—ask, adapt, author, and analyze—and it describes family activities, literacy questions, and design dimensions for each of the dimensions. Although Touretzky et al. propose five big main ideas that children should learn about AI technologies (Touretzky et al., 2019) in their framework, our framework

Table 10.1
The 4 As: proposed framework for families' AI literacy dimensions

AI literacy layer	Family activity	AI literacy question	AI design guideline
Ask	Interact fluently with an existing AI application or technology	How do you make it do . . . ? Do you . . . ? Are you . . . ?	Transparency Explainability
Adapt	Modify or customize an AI application to serve their needs	How do I modify it?	Personalization Transparency
Author	Create a new AI application	How do I make a new one?	Progressive Disclosure
Analyze	Analyze the data and the architecture of their AI application and modify it to test different hypotheses	How does it work? What if . . . ?	Systemic Reframing

focuses on children as active learners and agents of change who can decide how AI should work, not just discover its current functionalities. Another contribution of our framework is that it also addresses parents and tries to engage and support them in making more informed and meaningful use of the smart devices they might integrate into their homes.

Kids and Parents Ask AI

In prior studies, we investigated the challenges and opportunities of children growing up with digital technologies and their impact on the digital divide. In this context, access to AI literacy for families could prevent an AI divide for the generations of children growing up with smart technologies. With intelligent agents in the home, children do not need to read and write to access the internet; they can ask an agent any question or request, and the device will return the first result with a humanlike voice and friendly prosody. What seems at first to be a playful interaction between a child and a voice assistant can easily trigger events of real consequences (stories of children buying dollhouses and candy with Alexa without parental approval has already made national news). Our prior work (Druga et al., 2017) shows that overall, children found the AI agents to be friendly and trustworthy but that age strongly affected how they attributed intelligence to these devices. Younger participants (4–6 years old) were more skeptical of the devices' intelligence, while most older children (7–10 years old) declared the devices were more intelligent than they were. In a preliminary study, we found that older children mirrored their parents' choices for the smarter agent and used very similar explanations and attributions, even if they participated in the study independently (Beneteau et al., 2020; Druga et al., 2018). These findings build on work in developmental and early cognitive psychology (Gopnik, 2020) to underline the importance of leveraging children's natural tendency to "think like a scientist" when interacting with smart technologies.

Families Adapt AI

To compare how children use VUIs in different countries, we studied 102 children (7–12 years old) from four different countries (US, Germany, Denmark, and Sweden). The way children collaborated and communicated while describing their AI perceptions and expectations were influenced by both their socioeconomic and sociocultural background. Children in low- and medium-SES schools and community centers were better at collaborating compared to children in high-SES schools. However, children in low- and medium-SES centers had a harder time advancing because they had less experience with coding and interacting with these technologies. Our findings show that children outside the US were overall more critical and skeptical of the agent's intelligence and truthfulness (Anders, 2019; Druga et al., 2019) and had less exposure to these technologies.

Author AI: From Coding to Teaching Machines

Today, children cannot easily design their own AI devices, program their connected toys, or teach them proper behavior. However, some initiatives have started to design tools and platforms that enable youth to author with AI (Code.org, n.d.-a; Druga, 2018; "A guide to AI extensions to Snap!," n.d.; "Machine Learning for Kids," n.d.).

STEAM education has become a priority for schools and families around the world, and initiatives like Hour of Code and Scratch Days are currently reaching tens of millions of students in 180-plus countries (Code.org, n.d.-b). Learning how to program is also integrated into the curriculum in high schools across the UK and US. Meanwhile, parents are investing more resources to get their children involved in local technology and science clubs, camps, and coding events. Most of the educators, parents, and policy-makers are starting to recognize programming as a new literacy, which enables our youth to acquire and apply computational thinking skills. The technology used at home and in the classroom is changing fast. These advancements raise the opportunity not only to teach children how to code but also how to teach computers and embodied agents by training their own AI models or using existing cognitive services (Druga, 2018). An example of these kinds of AI coding platforms is shown in figure 10.2.

In a series of longitudinal studies, we found that programming and training smart devices changes the way children attribute intelligence and trust to smart devices. Participants from various SES backgrounds and different learning settings (public schools, private schools, community centers) became significantly more skeptical of AI's smarts once they understood how the AI worked (Druga, 2018; Druga et al., 2019). In traditional coding, children are used to sending a series of instructions to a machine and seeing how the code is compiled and executed. In AI learning, students have to understand the role of data and how it might influence the way machines execute algorithms (Cassell et al., 2000; Mioduser & Levy, 2010). Mioduser and Levy (2010) explored how children could understand robots' emergent behavior by gradually modifying the robots' environment. They discovered that young people are capable of developing a new schema when they can physically test and debug their

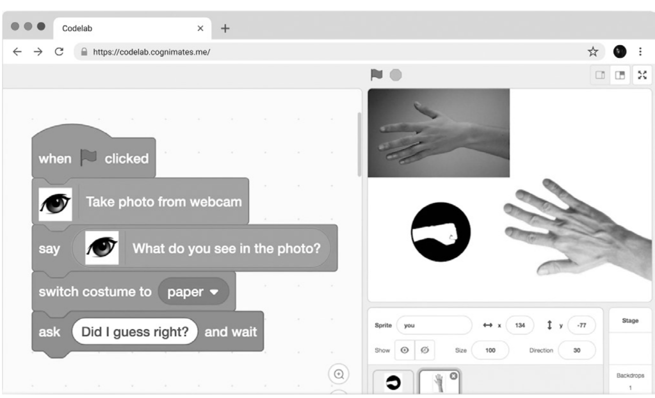


Figure 10.2
Examples of AI coding platforms (BlockStudio and Cognimates) piloted with families during our study.

assumptions. They also showed that the number of rules and new behaviors should be introduced gradually in the coding activity.

Programmability Helps to Analyze AI

Prior human-computer interaction (HCI) studies analyzing adults' mental models of AI technologies found that even a short tutorial with a researcher (i.e., 15 minutes) can significantly increase the soundness of participants' mental models. This phenomenon was consistent in Kulesza et al.'s study on intelligent music recommender systems and Bansal et al.'s study on the effect of different kinds of AI errors (Bansal et al., 2019; Kulesza et al., 2012). More so than users' explicit mental models, research on AI systems in HCI has focused on explainability and trust. Rutjes et al. (2019) argue for capturing a user's mental model and using it while generating explanations. At the same time, Miller (2019) invoked the concept of mental models through ideas of reconciling contradictions and our desire to create shared meaning in his comprehensive review of social science related to explainable AI.

When trying to understand how children and families analyze AI, we notice that programmability can play a significant role in influencing children's perception of smart agents' intelligence (Duuren, 1998; Scaife & Duuren, 1995; Scaife & Rogers, 1999). Additionally, parental mental models and attitudes can also influence how the children attribute intelligence to smart devices (Druga et al., 2018). Within this frame, we define sensemaking as a process by which people come across unfamiliar situations or contexts but need to process and understand to move forward (Klein et al., 2006). By creating activities and technologies that help families generate and test various hypotheses about how smart technologies work, we allow family members to not only test and understand how AI works; we also allow them to engage in systematic reframing and imagine how AI should work in order to support meaningful family activities (Dellermann et al., 2019).

The 4A Framework in Action

Ask dimension: Identify AI bias When we initially asked children to describe what bias means and give examples of bias as part of the co-design sessions (see figure 10.3), we found ourselves at a crossroads as we realized none of our participants understood what this term means. We quickly noticed, however, that children understood the notions of discrimination and preferential treatment and knew how to identify when technology was treating specific groups of people unfairly.

"Bias? It means bias," said L, a 7-year-old boy. During the initial discussion in the first study session, we tried to identify examples of bias that children could relate to, such as cookies or pet preferences. When talking about cat people versus dog people, D, a 9-year-old girl, said, "Everything they own is a cat! Cat's food, cat's wall, and cat. . . ." We then asked kids to describe dog people. A, an 8-year-old boy, answered: "Everything is a dog! The house is shaped like a dog, bed shaped like a dog." After children shared these two perspectives, we discussed again the concept of bias referring to the assumptions they made about cat and dog people. A summary of

Figure 10.3
Examples of families engaging with the smart toys activity during our co-design sessions.

the types of bias identified by children in sessions one and two is shown in figure 10.4.

Race and ethnicity bias: In the final discussion of the first session, children were able to connect their examples from daily life with the algorithmic justice videos they had just watched. “It is about a camera lens which cannot detect people in dark skin,” said A, while referring to other biased examples. We asked A why he thought the camera failed in this way, and he answered: “It could see this face, but it could not see that face . . . until she puts on the mask.” B, an 11-year-old girl, added, “It can only recognize White people.” These initial observations from the video discussions were later reflected in the children’s drawings. When drawing how the devices work, some children depicted how smart assistants separate people based on race. “Bias is making voice assistants horrible; they only see White people,” said A in a later session while interacting with smart devices.

Age bias: When children watched the video of a little girl having trouble communicating with a voice assistant because she could not pronounce the *wake* word correctly, they were quick to notice the

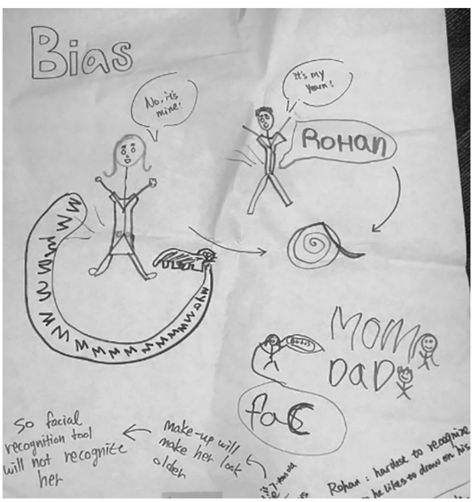


Figure 10.4
Examples of bias identified by children in sessions one and two.

age bias. "Alexa cannot understand baby's command because she said *Lexa*," said M, a 7-year-old girl. "When I was young, I did not know how to pronounce *Google*," she added, empathizing with the little girl in the video. Another boy, A, jumped in, saying: "Maybe it could only hear different kinds of voices," and shared that he does not know Alexa well because "it only talks to my dad." Other kids agreed that adults use voice assistants more.

Gender bias: After watching the video of the gender-neutral assistant and interacting with the voice assistants we had in the space, M asked: "Why do AI all sound like girls?" She then concluded that "mini Alexa has a girl inside, and home Alexa has a boy inside," and said of the mini Alexa: "I think she is just a copy of me!" While many of the girls were not happy that all voice assistants have female voices, they recognized that "the voice of a *neutral-gender* voice assistant does not sound right" (B, 11 years old). These findings are consistent with the UNESCO report on implications of gendering the voice assistants, which shows that having female voices for voice assistants by default is a way to reflect, reinforce, and spread gender bias (UNESCO, EQUALS Skills Coalition, 2019).

Adapt dimension: Trick the AI In the second design session, we invited participants to engage directly with the smart technologies and see if they could trick them. We wanted to provide the children with concrete ways in which they could test the device's limitations and bias, and we learned from our prior studies that children enjoy finding glitches and ways to make a program or a device fail (Druga, 2018). Such prompts not only give them a sense of agency but also provide valuable opportunities for debugging and for them to test their hypotheses about how the technology works. During our workshop, children imagined and tested various scenarios for tricking the different smart devices and algorithmic prediction systems. When playing with Anki's social robot, Cozmo, they decided to disguise themselves with makeup, masks, glasses, or other props so the robot could not recognize them anymore. They also decided to disguise other robots as humans to see whether doing so would trick the robots' computer vision algorithm. Children also used this strategy in our prior AI literacy workshops for families in Germany, and it is a fun activity that could easily be replicated at home.

When playing with the Quick, Draw! app, children were at first amazed at how quick and efficient the program was in guessing their drawings, so they decided to deploy many strategies to confuse the program. They first tried to draw nonsensical drawings to see if they would still get object predictions. They then decided that multiple children should try to draw on the same device at the same time so that the program would have a hard time keeping up with their drawing speed. When interacting with Alexa, the children probed it in various ways to find out whether it was biased. For example, they tried to speak Spanish to see if the device would recognize a new language; they used different names for calling the device *Lexa* to see if it could interact with more informal language; they asked "silly" questions to see if the device could engage in child play (e.g., "Call me 'princess'"); and they also tried to see if it could sing songs from different locations, such as the North Pole or the Indian Ocean. Very often, children built on each other's questions during the interaction and helped each other reformulate a question when needed. This finding is consistent with prior work in this field that demonstrates how much peers or family members can help repair communication breakdowns when interacting with voice assistants (Beneteau et al., 2019; Druga et al., 2017). While trying to probe and trick the voice assistant, children voiced several privacy concerns. "Amazon can hear everything users have said to their Alexas," said A, who then added, "Alexa buys data, takes data, and gives it to people who build Alexa." D was worried that "the tiny dots on Alexa are tiny eyes where people can see users," so she decided to cover the device with Post-it notes. From these examples, we see how children's privacy concerns can vary widely based on their naive theories (Inagaki, 1993), experiences with these technologies, and conversations they had had with or heard from their parents.

Author dimension: Design, code, teach the AI The democratization of current AI technologies allows children to communicate with machines not only via code but also via natural language and computer vision technologies. These new interfaces make it easier for children to control and even "program" an agent via voice, but they make it harder for children to debug the machine when it does not behave the way they expect. During our design sessions, children had the opportunity to discover a series of AI programming applications individually before using them with their parents. Sometimes families would start by playing with example games (figures 10.5a and 10.5b) that would recognize their gestures or objects. We would then ask them to make the games more or less intelligent. Other times families would come up with their project ideas and would start a program from scratch. We would ask the children to explain specific concepts from their project. For example, one of the researchers asked a child, M, "What does the loop mean?" M answered by drawing a circle in the air. We also asked both children and parents to reflect on how they could make the technology suitable and meaningful for their families. D's older sister said they could program the Sphero ball robot for "maybe dog chasing."

In all the authoring activities, families were trying to test their programs in various ways, moving their bodies together, standing up and sitting down. Meanwhile, one of the family members was going back and forth to modify the code blocks or the parameters of the smart games to see what would happen. Children and parents engaged in a balanced partnership, especially when using the applications where it was straightforward for multiple people to take turns interacting with the program (i.e., Quick, Draw!, Cognimates motion games, Teachable Machine vision training). Similar to prior studies, parents helped scaffold their children's behavior when interacting with robots or interactive devices together (Chang & Breazeal, 2011; Freed, 2012).

When M and her dad were playing together with the Teachable Machine platform (see figure 10.5), the dad would frequently probe his daughter with helping questions. For example: “So I put in 150 pictures, and you put in 25, so that model knows me better because I put more pictures in it. The more pictures I put in, the more the model will learn. How would you fix it?” he asked. M replied, “Add

Figure 10.5
Examples of children coding a game with BlockStudio and a family training a custom model with Teachable Machine.

more," and proceeded to add more pictures of herself. When she realized she could not add more pictures after a model was trained, she would say, "No, we have to redo it. Daddy goes first this time." After training their model for a second time, M and her dad tried to trick it, and both faced the camera at the same time to see which one would be recognized. M noted that they looked very similar to the machine but that because she had a pink bow, she thought the machine could recognize her. She thought of another way to trick the machine by giving her pink bow to her dad.

We observed the same behavior when families interacted with voice assistants. All family members helped each other repair various communication breakdowns, as in prior studies (Beneteau et al., 2019). For example, R's dad was trying to get the voice assistant to act like a cat by saying "meow" when talking to the device. "Oh, you have to say something," replied R, his 11-year-old son, who then added, "If you wanna wake her up, you should say something like *Alexa*." At his command, the device turned blue, and R said, "Meow." After, the voice assistant started to meow.

From these examples, we see how children build on experiences and skills developed in prior study sessions for probing the technology as they are designing it, either by asking it questions, trying to trick its games, debugging collaboratively with their families, or teaching and supporting each other. In this way, our ask, adapt, and author framework dimensions become intertwined in practice, helping families better understand and control AI technologies.

Analyze dimension: How does it work? How do we make it better? The last step in our design sessions with families was critically analyzing the technologies discussed, used, or created in all the other study sessions. This critical analysis was done in a group discussion at the end of the study, in which children, parents, and researchers participated in a circle. The analysis was also done throughout the other sessions every time we asked participants to draw and explain how the devices worked and what they had inside. With these prompts, we aimed to discover the families' mental models of AI technologies and observed how these explanations drew on or influenced their direct interaction with smart devices. The analyze discussion also elicited systematic reframing so that families could reflect on how they might use AI systems better in the future and to think about when and if they should use such technologies.

What is inside? To help uncover how children conceptualize smart devices, we asked them to draw what was inside the device and explain how it worked. Children resorted to various representations and explanations: a computer, a series of apps, a robot, a phone, or a search engine was inside the device. "There is a search engine inside the Alexa, but I do not know what it looks like," said L, a 10-year-old boy.

Y and S, two 9-year-old girls, said that there was an army of people who sit at their computers inside the "Company of Alexa" and reply to all the questions after they research the answers online. "There is a bunch of cords and a speaker inside the Alexa. It would connect to a computer and link it to Amazon people. If the question is 'What is the weather?' it [the person] would search the weather and type it up and let Alexa say it," said Y, a 9-year-old girl.

The most common analogy children made was of the mobile apps they are so familiar with. Children imagined how the voice assistant would use different mobile apps depending on the question the user asks. D, another 9-year-old girl, also imagined how the different devices were linked to each other: "If Alexa does not know an answer, it asks other Alexa[s] first before asking Amazon. Once one Alexa gets the answers . . . every single Alexa in the world will get that answers." The younger children (6–7 years old) provided more vitalistic explanations, consistent with prior studies (Inagaki, 1993). "There is a brain inside Alexa, and there is a part that connects to a computer with a speaker. The speaker will shout out the answer," said M, a 7-year-old girl. The older children (8–11 years old) had a very different explanation, primarily related to other technologies or applications they were using: "Alexa looks at every place it can search for an answer: Amazon, YouTube, internet, weather, map, anyplace," said A, an 8-year-old boy. "The database is a box with stuffs in it. The stuffs are statements you tell Alexa," added R, an 11-year-old boy.

It is as simple as 2+2: During the design sessions, children tried to validate their mental models by probing the different devices with questions. Children also tried to find out the age of the devices to determine how much they could trust them. Children were disappointed by the answer Alexa gave them when they asked how old it was: "It is as simple as 2+2." They described this answer as "questionable," as they found it hard to believe a voice assistant could possess so much information at the age of 4. B said the assistant must be at least 20 years old.

When children would find bugs or limitations in the device's answers, they thought the errors happened because the device "relies too much on the internet." Children requested to know who programmed the voice assistant so they could understand why the device was lying about its age. From this example, we see how our participants were able to draw on prior workshop experiences, not only understanding how the device behavior was linked to the way it was programmed but also figuring out what questions to ask in order to test the device.

Discussion

Our modern world is governed by the decisions made through AI and algorithms. While these tools show incredible promise in health care, education, and other fields, they also need to support ways in which people (mainly from vulnerable and marginalized populations) can carefully critique how AI could amplify racism, sexism, and other forms of discrimination. For people to start considering algorithmic justice early in life, we must find ways they can develop forms of literacy around AI. We argue that AI justice and AI literacy begins in early interactions, inquiries, and investigations in the family.

AI literacy, however, is not a form of knowledge that can be simply taught in a didactic and lecture-based form (Druga, 2018). Instead, designers need to consider how to promote sensemaking, collaboration, questioning, and critical thinking. How can they design future AI systems for families that tap into the idea of "children as scientists" and leverage children's curiosity and both the explore and exploit paradigms? Prior work shows that children are developmentally primed for this type of exploration (Gopnik, 2020), and we believe it is a missed opportunity to not provide AI literacy opportunities by designing future smart technologies and via parenting.

Based on our prior research and this study's findings, we propose a novel AI literacy framework for designers and educators to consider in order to support families' critical understanding and use of AI systems. We believe it is important to consider this design framework in the context of our current analysis of nested ecological systems (Bronfenbrenner, 1994).

In asking sessions, children and families can inquire and interact with AI agents through various means, such as calling out with voice interactions, drawing, and playing. However, embedded in these interactions with asking are privacy policies that need to be transparent for families (exosystem). Families have several questions about the impact and interplay of privacy, technology, policy, and their children (Zeng et al., 2017). Therefore, how do we support families to ask and interact with AI agents in a way that deems their information safe and confidential? Designers also need to consider how at-home interactions happen between children and families (microsystems). In this context, are families able to collaborate and ask AI agents together? How do prior relationships with technology in families mediate how comfortable family members are engaging with AI at home?

With adaptation sessions, families are shifting and mitigating their perceptions and engagements around AI to fit their contexts. However, as families adapt to AI, questions of negotiation and power remain (Barocas & Selbst, 2016). AI systems are unable to code switch and recognize children and adults (Beneteau et al., 2019), raising the risk that age-inappropriate content may be accessed by children. How does AI think about more substantial cultural capital and social contexts (macrosystems) of families? For instance, bilingual families can switch and merge languages (e.g., Spanglish) in their routine conversations with one another. For AI voice assistants, this means having to adopt a single language. Similarly, AI systems have difficulty recognizing different languages and accents (macrosystems). In this case, families who may have grown together in specific social and cultural norms now face systems that are unable to adapt to these larger macrosystems.

For the author dimension, families need a chance to build and create in order to develop AI literacy. We ask, though, who has an opportunity to build? Even if designers create authoring systems for AI engagement, those systems can depend solely on technology infrastructure at home (exosystems) (Riddlesden & Singleton, 2014). Authoring may also mean learning how to build, which may privilege individual families in communities, libraries, schools, and networks that can teach and build knowledge capacity.

Finally, under analyze, AI learning tools can be designed with collaboration and sensemaking in mind (Ash, 2004; Paul & Reddy, 2010). This approach assumes that different family units work together (microsystem). Therefore, how is a careful reflection on AI designed to deal with real family constraints, like working families, families with limited time, and families who always move (i.e., children living between households)? How might designers create activities and technologies that support diverse families, allowing those families to generate and test various hypotheses about how smart technologies work and systematically reframe how AI should work to support meaningful and inclusive family activities (Dellermann et al., 2019)?

Overall, while complex ecological systems need to be considered within design frameworks, there are still takeaways for families who have adopted AI literacy and justice. Our study shows that with the ask, adapt, author, and analyze dimensions, parental roles and relationships still matter when families are learning about AI together. Aarsand (2007) describes “asymmetrical relations” between parents and children as both a challenge and an opportunity for families to jointly engage with assumptions about media like computers and video games. The “digital divide”—through which children are considered experts with digital media while adults are positioned as novices—becomes a “resource for both children and adults to enter and sustain participation in activities” (Aarsand, 2007). Children can teach parents about AI technologies, but it is also the parents’ responsibility to teach children about the values in their community that matter and how AI tools and systems align with these values (Friedman et al., 2008).

Design Features That Encourage AI Literacy for Families

Using our findings, we can examine the conditions and processes that our family AI literacy framework could support. We use our findings to show how the ask, adapt, author, and analyze dimensions can lead families to adopt a critical understanding of AI (Druga, 2018; Druga et al., 2019), specifically through a balanced engagement with these new technologies (Sobel et al., 2004; Takeuchi & Stevens, 2011; Yip et al., 2017). This balanced engagement involves:

- **Mutual engagement** (i.e., multiple family members should be equally motivated to participate): Families in this study were able to participate in different ways, whether they were asking several questions to voice assistants, playing and authoring together with new AI systems, or trying to analyze how bias is introduced into smart technologies.

- **Dialogic inquiry** (i.e., inquiry by families inspires collaboration and meaning-making): Families can try to analyze the AI systems and try to figure out how they work. They can also determine how the AI systems need to adapt to their families' culture, rules, and background.
- **Co-creation** (i.e., people create shared understanding through co-usage): Parents and children can come together to **ask**, **adapt**, **author**, and **analyze** AI systems in order to find out what they know and what they would like to know more about.
- **Boundary crossing** (i.e., AI spans time and space): Families can consider how AI systems are pervasive in multiple technologies, whether in internet searches, YouTube recommendation systems, or voice assistants of multiple forms. If families can recognize how pervasive AI is becoming on many platforms, they can shape how AI itself is crossing boundaries.
- **Intention to develop** (i.e., families gain experience and development): Families can consider how they are adapting to AI systems. For instance, are the questions they are asking voice assistants changing? Are families noticing when AI systems may be present? Interestingly, families can develop as they understand how AI systems themselves are adapting to different people and contexts.
- **A focus on content, not control** (i.e., interface does not distract from interaction): With some AI systems, families can engage in multiple straightforward ways. Through asking voice assistants questions, seeing if AI systems can recognize drawings and sketches, and engaging with computer vision models, families can now question and critique AI systems using many simple mechanics.

Conclusion

Our aim in designing technologies is to ensure we are supporting families in raising a generation of children who are not merely passive consumers of AI technologies but rather active creators and shapers of its future. With our AI literacy framework, we aim to encourage and enable families to learn how to develop a critical understanding of AI. We propose this framework from an ecological systems theory perspective and present examples of implications for supporting family AI literacy across various nested layers of our society. As designers of technologies, we aim to support a diverse population of children and adults and provide significant inspiration and guidance for future designs of more inclusive human-machine interactions. We hope that by democratizing access to AI literacy through tinkering and play, we will enable families to step in and decide when and how they wish to invite AI into their homes and lives.

References

Aarsand, P. A. (2007). Computer and video games in family life: The digital divide as a resource in intergenerational interactions. *Childhood*, *14*(2), 235–256.

Anders, L. (2019). *AI for kids—it is our responsibility to enable children worldwide to engage with artificial intelligence*. Medium. Retrieved April 11, 2020, from https://medium.com/@_tlabs/ai-for-kids-it-is-our-responsibility-to-enable-children-worldwide-to-engage-with-artificial-ec0d5c627945

Anderson, M. (2019). *Mobile technology and home broadband 2019*. Pew Research Center. https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/06/PI_2019.06.13_Mobile-Technology-and-Home-Broadband_FINAL2.pdf

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Ash, D. (2004). Reflective scientific sense-making dialogue in two languages: The science in the dialogue and the dialogue in the science. *Science Education*, *88*(6), 855–884.

Banerjee, R., Liu, L., Sobel, K., Pitt, C., Lee, K. J., Wang, M., Chen, S., Davison, L., Yip, J. C., Ko, A. J., & Popović, Z. (2018). Empowering families facing English literacy challenges to jointly engage in computer programming. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 622.

Bansal, G., Nushi, B., Kamar, E., Weld, D. S., Lasecki, W. S., & Horvitz, E. (2019). Updates in human-AI teams: Understanding and addressing the performance/compatibility tradeoff. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*, 2429–2437.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, *104*(3), 671.

Barron, B. (2004). Learning ecologies for technological fluency: Gender and experience differences. *Journal of Educational Computing Research*, *31*(1), 1–36.

Beneteau, E., Guan, Y., Richards, O. K., Zhang, M. R., Kientz, J. A., Yip, J., & Hiniker, A. (2020). Assumptions checked: How families learn about and use the Echo Dot. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(1), Article 3, 1–23.

Beneteau, E., Richards, O. K., Zhang, M., Kientz, J. A., Yip, J., & Hiniker, A. (2019). Communication breakdowns between families and Alexa. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 243.

Berthelsen, D., & Walker, S. Parents' involvement in their children's education. *Family Matters*, 2008(79), 34.

Bronfenbrenner, U. (1994). Ecological models of human development. *Readings on the Development of Children*, *2*(1), 37–43.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Conference on Fairness, Accountability and Transparency*, 77–91.

Cassell, J., Sullivan, J., Churchill, E., & Prevost, S. (2000). *Embodied conversational agents*. The MIT Press.

Chang, A., & Breazeal, C. (2011). Tinkrbook: Shared reading interfaces for storytelling. *Proceedings of the 10th International Conference on Interaction Design and Children*, 145–148.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.

Chklovski, T., Jung, R., Fofang, J. B., Gonzales, P., Hub, B. T., & La Paz, B. (2019). Implementing a 15-week AI-education program with underresourced families across 13 global communities. *International Joint Conference on Artificial Intelligence*.

Code.org. (n.d.-a). Retrieved April 11, 2020, from https://code.org/

Code.org. (n.d.-b). *Statistics*. Retrieved April 11, 2020, from https://code.org/promote

Cole, M., Amsel, E., & Renninger, K. (1997). Cultural mechanisms of cognitive development. In E. Amsel, K. A. Renninger, & A. Renninger (Eds.), *Change and Development: Issues of Theory, Method, and Application* (pp. 245–263). Routledge.

Coyne, S. M., Radesky, J., Collier, K. M., Gentile, D. A., Linder, J. R., Nathanson, A. I., Rasmussen, E. E., Reich, S. M., & Rogers, J. (2017). Parenting and digital media. *Pediatrics*, *140*(Suppl. 2), S112–S116.

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., & Ebel, P. (2019). The future of human-AI collaboration: A taxonomy of design knowledge for hybrid intelligence systems. *Proceedings of the 52nd Hawaii International Conference on System Sciences, 274–283*. doi:10.24251/HICSS.2019.034

DiSalvo, B., Khanipour Roshan, P., & Morrison, B. (2016). Information seeking practices of parents: Exploring skills, face threats and social networks. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 623–634.

Druga, S. (2018). *Growing up with AI: Cognimates: From coding to teaching machines* [Doctoral dissertation, Massachusetts Institute of Technology]. DSpace@MIT.

Druga, S., Vu, S. T., Likhith, E., & Qiu, T. (2019). Inclusive AI literacy for kids around the world. *Proceedings of FabLearn 2019*, 104–111.

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (2017). Hey Google is it OK if I eat you? Initial explorations in child-agent interaction. *Proceedings of the 2017 Conference on Interaction Design and Children*, 595–600.

Druga, S., Williams, R., Park, H. W., & Breazeal, C. (2018). How smart are the smart toys? Children and parents' agent interaction and intelligence attribution. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 231–240. https://doi.org/10.1145/3202185.3202741

Druin, A. (2002). The role of children in the design of new technology. *Behaviour and Information Technology (BIT)*, *21*(1), 1–25.

Duuren, M. V. (1998). Gauging children's understanding of artificially intelligent objects: A presentation of counterfactuals. *International Journal of Behavioral Development*, *22*(4), 871–889.

European Commission. (2011). *Special Eurobarometer 359: Attitudes on data protection and electronic identity in the European Union*. Conducted by TNS Opinion & Social at the request of Directorate-General Justice, Information Society & Media and Joint Research Centre. Directorate-General for Communication

Federal Trade Commission. (2017). *Children's online privacy protection rule: A six-step compliance plan for your business*. https://www.ftc.gov/business-guidance/resources/childrens-online-privacy-protection-rule-six-step-compliance-plan-your-business

Ferguson, A. G. (2012). Predictive policing and reasonable suspicion. *Emory LJ*, *62*, 259.

Freed, N. A. (2012). *"This is the fluffy robot that only speaks French": Language use between preschoolers, their families, and a social robot while sharing virtual toys* [Doctoral dissertation, Massachusetts Institute of Technology]. DSpace@MIT.

Friedman, B., Kahn, P. H., & Borning, A. (2008). Value sensitive design and information systems. In K. E. Himma & H. T. Tavani (Eds.), *The Handbook of Information and Computer Ethics* (pp. 69–101).

Gebru, T. (2019). *Oxford handbook on AI ethics book chapter on race and gender*. https://arxiv.org/abs/1908.06165

Gonzales, A. (2017). Technology maintenance: A new frame for studying poverty and marginalization. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 289–294.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B*, *375*(1803). https://doi.org/10.1098/rstb.2019.0502

Grossman, J., Lin, Z., Sheng, H., Wei, J. T.-Z., Williams, J. J., & Goel, S. (2019). *Mathbot: Transforming online resources for learning math into conversational interactions*. Association for the Advancement of Artificial Intelligence.

Guha, M. L., Druin, A., Chipman, G., Fails, J. A., Simms, S., & Farber, A. (2004). Mixing ideas: A new technique for working with young children as design partners. *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community*, 35–42.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29.

Inagaki, K. (1993). Young children's differentiation of plants from nonliving things in terms of growth. *Biennial meeting of Society for Research in Child Development, New Orleans*.

Ito, M., Baumer, S., Bittanti, M., boyd, d., Cody, R., Stephenson, B. H., Horst, H. A., Lange, P. G., Mahendran, D., Martınez, K. Z., Pascoe, C. J., Perkel, D., Robinson, L., Sims, C., & Tripp, L. (2009). *Hanging out, messing around, and geeking out: Kids living and learning with new media*. The MIT Press.

Khan, K. (n.d.). A guide to AI extensions to snap! [(Accessed on 04/11/2020)]. https://ecraft2learn.github.io/ai/

Kahn, P. H., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Ruckert, J. H., & Gill, B. (2011). The new ontological category hypothesis in human-robot interaction. *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 159–160.

Kahn, Jr., P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, J. H., & Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology*, *48*(2), 303.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, *21*(5), 88–92.

Kulesza, T., Stumpf, S., Burnett, M., & Kwan, I. (2012). Tell me more? The effects of mental model soundness on personalizing an intelligent agent. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–10.

Kulick, D., & Stroud, C. (1993). Conceptions and uses of literacy in a Papua New Guinean village. In B. Street (Ed.), *Cross-Cultural Approaches to Literacy* (pp. 30–61). Cambridge University Press.

Lau, J., Zimmerman, B., & Schaub, F. (2018). Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1–31.

Lee, M. K., Davidoff, S., Zimmerman, J., & Dey, A. (2006). Smart homes, families, and control. *Conference of the Design & Emotion Society*.

Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, *17*(4), 391–444.

Lievens, E. (2017, January 25–27). *Children and the GDPR: A quest for clarification and the integration of child rights considerations* (Panel: Generation Zero - Data & Digital Marketing Protections for Children and Teens under the GDPR, COPPA and the new FCC Privacy Rules) [Paper presentation]. Computers, Privacy & Data Protection: The Age of Intelligent Machines, Brussels, Belgium. http://hdl.handle.net/1854/LU-8505935

Livingstone, S. (2018). Children: A special case for privacy? *Intermedia*, *46*(2), 18–23.

Livingstone, S., Haddon, L., Görzig, A., & Olafsson, K. (2011). *Risks and safety on the internet: The perspective of European children: Full findings and policy implications from the EU kids online survey of 9–16 year olds and their parents in 25 countries*. EU Kids Online Network. http://eprints.lse.ac.uk/43731/1/Risks%20and%20safety%20for%20children%20on%20the%20internet%28lsero%29.pdf

Livingstone, S., Stoilova, M., & Nandagiri, R. (2019). *Children's data and privacy online: Growing up in a digital age: An evidence review*. London School of Economics and Political Science.

Lovato, S., & Piper, A. M. (2015). Siri, is this you? Understanding young children's interactions with voice input systems. *Proceedings of the 14th International Conference on Interaction Design and Children*, 335–338.

Machine learning for kids. (n.d.). Retrieved April 11, 2020, from https://machinelearningforkids.co.uk

McReynolds, E., Hubbard, S., Lau, T., Saraf, A., Cakmak, M., & Roesner, F. (2017). Toys that listen: A study of parents, children, and internet-connected toys. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5197–5207.

Medin, D. L., & Bang, M. (2014). *Who's asking? Native science, Western science, and science education*. The MIT Press.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38. https://www.sciencedirect.com/science/article/abs/pii/S0004370218305988?via%3Dihub

Mioduser, D., & Levy, S. T. (2010). Making sense by building sense: Kindergarten children's construction and understanding of adaptive robot behaviors. *International Journal of Computers for Mathematical Learning*, *15*(2), 99–127.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Paul, S. A., & Reddy, M. C. (2010). Understanding together: Sensemaking in collaborative information seeking. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, 321–330.

Pina, L. R., Gonzalez, C., Nieto, C., Roldan, W., Onofre, E., & Yip, J. C. (2018). How Latino children in the U.S. engage in collaborative online information problem solving with their families. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 140.

Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018). Voice interfaces in everyday life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 640.

Regulation (EU) 2016/679 of the European Parliament, Art. 8GDPR: Conditions applicable to child's consent in relation to information society services. https://gdpr-info.eu/art-8-gdpr/

Riddlesden, D., & Singleton, A. D. (2014). Broadband speed equity: A new digital divide? *Applied Geography*, *52*, 25–33. https://doi.org/10.1016/j.apgeog.2014.04.008

Rogers, J. (2019). Privacy included: Rethinking the smart home. *Internet Health Report*, Special Edition. https://foundation.mozilla.org/en/privacy-included/

Rogoff, B., Najafi, B., & Mejía-Arauz, R. (2014). Constellations of cultural practices across generations: Indigenous American heritage and learning by observing and pitching in. *Human Development*, *57*(2–3), 82–95.

Rosebery, A. S., Ogonowski, M., DiSchino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *The Journal of the Learning Sciences*, *19*(3), 322–357.

Ruan, S., He, J., Ying, R., Burkle, J., Hakim, D., Wang, A., Yin, Y., Zhou, L., Xu, Q., AbuHashem, A., Dietz, G., Murnane, E. L., Brunskill, E., & Landay, J. A. (2020). Supporting children's math learning with feedback-augmented narrative technology. *Proceedings of the Interaction Design and Children Conference*, 567–580. https://dl.acm.org/doi/abs/10.1145/3392063.3394400

Ruan, S., Jiang, L., Xu, J., Tham, B. J.-K., Qiu, Z., Zhu, Y., Murnane, E. L., Brunskill, E., & Landay, J. A. (2019). Quizbot: A dialogue-based adaptive learning system for factual knowledge. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13.

Rutjes, H., Willemsen, M., & IJsselsteijn, W. (2019). *Considerations on explainable AI and users' mental models*. Where is the Human? Bridging the Gap Between AI and HCI: Workshop at CHI'19, May 4-9, 2019, Glasgow, Scotland, UK. Association for Computing Machinery. http://www.martijnwillemsen.nl/recommenderlab/RutjesChHI2019ws.pdf

Scaife, M., & Duuren, M. (1995). Do computers have brains? What children believe about intelligent artifacts. *British Journal of Developmental Psychology*, *13*(4), 367–377.

Scaife, M., & Rogers, Y. (1999). Kids as informants: Telling us what we didn't know or confirming what we knew already. In A. Druin (Ed.), *The Design of Children's Technology* (pp. 27–50). Morgan Kaufmann Publishers.

Schön, D. A. (1987). *Educating the reflective practitioner*. Jossey-Bass Publishers.

Scribner, S., & Cole, M. (1981). Unpackaging literacy. In M. F. Whiteman (Ed.), *Writing: The Nature, Development, and Teaching of Written Communication: Volume 1: Variation in writing: Functional and linguistic-cultural differences*, 71–87.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, *28*(3), 303–333.

Stevens, R., & Penuel, W. R. (2010). Studying and fostering learning through joint media engagement. *Principal Investigators Meeting of the National Science Foundation's Science of Learning Centers*, 1–75.

Street, B. (2003). What's "new" in new literacy studies? Critical approaches to literacy in theory and practice. *Current Issues in Comparative Education*, *5*(2), 77–91.

Takeuchi, L., & Stevens, R. (2011). *The new coviewing: Designing for learning through joint media engagement*. The Joan Ganz Cooney Center at Sesame Workshop and

LIFE Center. [Report]. https://www.joanganzcooneycenter.org/wp-content/uploads/2011/12/jgc_coviewing_desktop.pdf

Technovation. (2019). *2019 Impact Report*. Retrieved April 11, 2020, from https://www.technovation.org/wp-content/uploads/2020/04/Technovation-2019-General-Impact-1.pdf

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K–12: What should every child know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*, 9795–9799.

UNESCO, EQUALS Skills Coalition. (2019). *I'd blush if I could: Closing gender divides in digital skills through education*. UNESCO Digital Library. Retrieved March 9, 2021, from https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=85

van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, *34*(4–5), 221–235.

Walsh, G., Druin, A., Guha, M. L., Foss, E., Golub, E., Hatley, L., Bonsignore, E., & Franckel, S. (2010). Layered elaboration: A new technique for codesign with children. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1237–1240.

Walsh, G., Foss, E., Yip, J., & Druin, A. (2013). FACIT PD: A framework for analysis and creation of intergenerational techniques for participatory design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2893–2902.

Walsh, G., & Wronsky, E. (2019). AI+ co-design: Developing a novel computer-supported approach to inclusive design. *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, 408–412.

Woodward, J., McFadden, Z., Shiver, N., Ben-hayon, A., Yip, J. C., & Anthony, L. (2018). Using co-design to examine how children conceptualize intelligent interfaces. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14.

Yardi, S., & Bruckman, A. (2012). Income, race, and class: Exploring socioeconomic differences in family technology use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3041–3050.

Yip, J. C., Sobel, K., Pitt, C., Lee, K. J., Chen, S., Nasu, K., & Pina, L. R. (2017). Examining adult-child interactions in intergenerational participatory design. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5742–5754.

Zeng, E., Mare, S., & Roesner, F. (2017). End user security and privacy concerns with smart homes. *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*, 65–80.

11 Do Educational Technologies Have Politics?

A Semiotic Analysis of the Discourse of Educational Technologies and Artificial Intelligence in Education

Paulo Blikstein and Izidoro Blikstein

Introduction

With every new technological generation, we seem to forget their politics. In his controversial “Do Artifacts Have Politics?,” Langdon Winner (1980) recounts roads designed to privilege private cars, cities planned to diffuse public protests, and industrial technologies intentionally implemented not to improve products but to displace workers’ unions. Winner’s claims and some of the historical facts he cites have been contested. However, his examples undoubtedly rekindle a decades-old debate about the relationships between designers’ intentions, their resulting technical or architectural artifacts, and the social systems in which those artifacts operate. Today, this conversation has moved from simply stating that designed artifacts might serve a political purpose to understanding the more nuanced ways in which larger systems enable buildings or inventions to “implement” their politics (see, e.g., Joerges, 1999; Latour, 2004; Woolgar & Cooper, 1999).

Understanding these systemic contexts is more important than ever as we strive to comprehend our modern world. While technology corporations operate under benign mottos such as “connecting people" or "organizing information," they do far more than connect and organize; they implement a vision for *reshaping* human interaction in line with their technologies and business models, displacing competing versions regardless of their quality. Moreover, these outcomes are only possible because these companies design artifacts and technologies that operate within systems that are particularly well suited to amplify their impact. Social media apps, for instance, would undoubtedly be less potent under a regulatory framework that limited their ability to collect data.

A societal awakening to the era of digital surveillance and manipulation (Zuboff, 2015) has led scholars to reexamine the *systemic* impact of modern digital technologies on all areas of human activity, including education. What was once believed to be unequivocally beneficial—such as universal, free access to educational materials—is now seen in light of many previous critiques of technology and society. More closely examining the politics of educational technologies and their enabling systems has become imperative. Foundational work in this field (e.g., Selwyn, 2010, 2013; Selwyn & Facer, 2013; Vakil, 2018; Watters, 2015a; Williamson, 2018a, 2018b) confirms that the issue is not just the creation of such learning technologies but also their (often malign) affinities with the larger sociotechnical systems that generated them. Going beyond the simplified critique about "artifacts having politics," we must examine how technological artifacts augment, enable, and facilitate specific, preexisting visions of education. What comes first: an education system that privileges testing and ranking, or apps that facilitate those operations—or do they coevolve?

Before proceeding with the remainder of this chapter, which builds on the aforementioned critiques by investigating the process by which the world became enamored with and extensively funded certain educational technologies, we must specify the "certain type" of educational technology we will discuss. Winner himself (2009) and others (e.g., Tyack & Cuban, 1995) often fail to understand that not all educational technologies are created equal and that ascribing unequivocal intentions to designers fails to capture the entire picture (Joerges, 1999): the politics of technological artifacts can go both ways. Yes, a computer can be used to mimic the traditional, oppressive classroom, but it can also offer students novel, subversive tools for knowledge creation to escape schoolified oppression (Buechley & Eisenberg, 2009; DiSalvo, 2014; Freire, 2014; Latour, 2004; Papert, 1980). Failing to understand the subversive "Papertian" or "Buechleyan" uses of computing, and denying students access to them, could well be another subtle form of oppression.

We first discuss "automated instructionist technologies," which we define as techniques and artifacts designed to teach students predetermined content through electronic media (such as video classes) accompanied by automated assessment and often aided by AI-powered systems. We also investigate the "enamoring" of the educational world with these technologies, despite abundant historical accounts of unfulfilled hyperbolic promises and decades of accumulating negative evidence on their efficacy (e.g., Cole et al., 2012; Ready et al., 2019; Reich, 2020; Watters, 2021). We hypothesize that this unparalleled feat of persistence was accomplished not only by sophisticated products or efficient marketing campaigns but also by the skillful construction of resilient discourse (Bakhtin, 1984).

As these discursive moves are subtle and barely visible to the untrained eye, we introduce in our next section the analytical tools of discourse analysis and semiotics, which we use to break the narrative into pieces and see what lies behind it. What we find is a standard and repeated three-step formula of controlling the narrative of what education is about, silencing dissenting voices such as those of experienced teachers, and defining—as a strawman—what counts as innovation. The implications for policy-makers, educators, and parents include realizing that automated technologies, rather than simply being a tool to improve schooling, have the potential to *change the nature of education itself* by molding school learning according to what the technology is capable of accomplishing.

Methods and Discourse Analysis

Semiotic analysis examines discourse for "barely perceptible" traces that might reveal the rationales and mental models that drive the message's creation. Thus, our first methodological move to get to the crucial presuppositions of the discourse of automated instructionist technologies is to *decipher not the visible but the intelligible*, in the same way that art appraisers distinguish authentic paintings from forgeries by inspecting seemingly insignificant details, or how Freudian psychotherapists dwell on minor lapses of memory or language (Ginzburg, 1991).

Our analytical lens also uses many of Bakhtin's discourse analysis theories. Our first tool, *dialogical discourse*, is a type of construction in which a product or idea is qualified based on the creation of an antagonist. As a result, people must inevitably consider other discourses that will be in dialogical opposition with their own (Bakhtin, 1984; Todorov, 1989). We also explore the idea of *polyphony*: discourse is not monophonic or autonomous but rather "spoken" by many voices intertwined in time and space. Finally, the analytic lens of *intertextuality* enables us to examine how authors—intentionally or not—use texts from the past and present to legitimize arguments in favor of their product or idea, requiring a semiotic archeological "excavation" to recover their significance. We use these three analytic lenses—intertextuality, polyphony, dialogism—to disentangle and reveal the ideas about education buried in the discourse of automated instructionist innovation. We do not mean that discourses are generated with the explicit goal of disguising the actual message but that they are long, dialogical, and polyphonic social constructions that might even escape the comprehension of their direct beneficiaries (Blikstein, 2020).

We also remember Benveniste's (1966) and Jakobson's (1968) note about discourse being more than the mere transmission of information. It also has a *connative function*, by which authors focus on generating a positive response and a favorable effect instead of concentrating only on the content of the message (a lesson well learned by contemporary politicians and marketers).

Our analysis forgoes the idea that a handful of technology entrepreneurs "disrupt" education or create their discourses in isolation. Conversely, intertextuality shows that these discourses are built collectively and in multilayered structures of validation, ideology, economic interests, and individual ambitions (Bakhtin et al., 1993). We also consider that these automated educational technologies are dissimilar from most other products marketed to students and parents (such as pens and notebooks), teachers (such as textbooks), or school districts (such as furniture). Unlike these more traditional products, educational technologies must do more than slightly surpass their competitors; they require a "transcendent" buy-in: they must be disruptive, innovative, revolutionary, and "game-changing." Thus, they have to be associated with powerful brands and personalities or connected to a grand, utopian educational progress theory. Otherwise, the *transcendent buy-in* is never achieved.

Automated educational technologies achieve this buy-in through advertising strategies that take the rules of engagement of traditional advertising to another level. In marketing a traditional product, say a chair, it is often not enough to say that a brand of chairs is better—you might also recruit orthopedists to record testimonies and encourage satisfied customers to post pictures on social media. Ultimately, this collection of "texts" from different sources not only speaks about the chair but also about the ethos of the company, its sustainable practices, and even the charitable acts of the CEO. Automated educational technologies employ an even more powerful and complex set of texts to eventually normalize and naturalize their message (Moles, 1958), using this polyphony and intertextuality to make the idea not just desirable but inevitable. We uncover these techniques and the messages they obscure by following the innovative and foundational methodologies and research put forth by the pioneering work of Audrey Watters (Watters, 2015b, 2021), who combines historiography, technical analysis, and critical theory.

Data Collection

Our methodology included three data collection phases. First, we collected the self-reported "mission statements" from the websites of 15 major edtech companies working on automated instructionist technologies (including AI in education, which is part of many of those technological solutions.) These companies were selected based on systematic online searches within specialized databases such as TechCrunch (https://techcrunch.com/tag/edtech/) and Crunchbase (https://www.crunchbase.com).

We then collected publicly available interviews with some of the prominent leaders in the field and news pieces in which they are quoted. These texts were then filtered for the most relevant content (i.e., by deleting redundant descriptions of the products, recountings of the companies' histories, or biographies of the CEOs), resulting in about 20,000 words of data spanning about eight years (from 2011 to 2019). The data were analyzed for themes and topics, and representative excerpts were selected.

Finally, we used web-scrubbing techniques to extract the most recent public news pieces with the keywords "AI in Education" and "MOOCs." These keywords do not represent the gamut of automated instructionist technologies available today, but they capture two main types of technologies from the past decade. We focused the search in this way to limit the number of results while capturing critical services within the world of automated instructionist technologies. This search resulted in 623 articles, from which we extracted the titles and first 20 words.

Our goal, however, is not to conduct a quantitative analysis of this data set. Our goal with the automated data collection was to capture a broad enough set of texts to use the lens of discourse analysis on a sample that would represent the industry's leading voices. Thus, we present our qualitative analysis of the first three data sources (mission statements, interviews, and news reports), using the web-scrubbed data only for triangulation: for every interview or excerpt selected for in-depth analysis, we would reexamine the larger data set to ascertain their typicality; if the excerpt seemed like an outlier or a statement that contradicted the typical ones in the data set, we instead chose a different excerpt for the qualitative analysis.

"Edtech": History and Discourses

To offer some context for the data analysis, we start with a brief historical narrative of attempts to bring technological artifacts into education, from Thorndike to the Silicon Valley, situating their ideas and inventions within a century-long tradition of automated educational technologies.

The Early Days and Pressey's Teaching Machine

Since the mid-nineteenth century, many inventions were supposed to revolutionize schooling, such as early slide projectors and erasable writing devices (see, e.g., Cuban, 1986). Watters's (2021) historical account on educational technologies describes a machine for teaching spelling from 1866, Thorndike's "personalized textbook" from 1912, Aikins's contraption from 1913 (an "educational appliance to teach any subject"), Pressey's testing machine from 1924, and an IBM test-scoring machine from 1937. Considering our analytical framework, we interpret more recent waves of innovation not as mere copies of previous generations' work but as an ongoing intertextual dialogue between multiple generations of educational technologists, by which ideas, rationales, and justifications flow through time, alternating between "hibernation" and intense public interest.

The Pressey machine is the most famous of the first wave of automated devices and one of the first cases of the phenomenon that we repeatedly explore: the creation of a dialogical antagonist (Bakhtin et al., 1993) that justifies a product's existence without detailing *how* the new product will fulfill its goals. The Pressey contraption auto-graded multiple-choice questions and, despite its simplicity, was touted as revolutionary: "There must be an industrial revolution in education, in which educational science and the ingenuity of educational technology combine to modernize the grossly inefficient and clumsy procedures of conventional education. [This revolution would free teachers to develop] in pupils fine enthusiasms, clear thinking, and high ideals" (Pressey, 1933, as cited in Watters, 2021).

Pressey's discursive antagonist is clear: the "grossly inefficient and clumsy procedures" of education. However, his solution, rather than some advanced device to bring about "clear thinking and high ideals," was simply a mechanical contraption for revealing the correct answer to a multiple-choice question. In reading his quote, however, it is hard to oppose his education diagnosis—who would advocate for the "clumsy procedures" of education? Here, Pressey seems to follow what Benveniste (1966) would formalize a few decades later: while his message is a gross overgeneralization and the content is imprecise and hyperbolic, it creates a positive effect of persuasion, which distracts us from questioning the connection between the stated problem and the proposed solution. It is not clear how a standardized testing machine would develop "high ideals"—but a positive reaction to the discourse is created nonetheless. The competent creation of an indefensible antagonist makes us want to jump on board with Pressey's invention—even if neither the machine's functioning nor its connection to the inventor's critique of traditional education are well defined. This strategy would be perfected by later generations of automated educational technologies before finding a permanent place in the field's playbook.

Skinner's Teaching Machine

Skinner continued Pressey's agenda in his teaching machine but brought along a more comprehensive theory (behaviorism) and a powerful brand (Harvard University), amplifying the impact and reach of this type of device. Notably, he refined his discourse, too. In a 1954 video recorded to introduce Skinner's invention, we see intertextuality at work:

> I should like to discuss some of the reasons why studying with the help of a teaching machine is often dramatically effective . . . as soon as the student has written his response, he operates the machine and learns immediately whether he is right or wrong. This is a great improvement over the system in which papers are corrected by a teacher where the student must wait perhaps till another day to learn whether or not what he has written is right. Such immediate knowledge has two principal effects: it leads most rapidly to the formation of correct behavior; the student quickly learns to be right. But there is also a motivating effect: the student is free of uncertainty or anxiety about his success or failure, his work is pleasurable, he does not have to force himself to study. A classroom in which machines are being used is usually the scene of intense concentration. (Skinner, 1954)

Skinner begins with similar discursive moves to Pressey. He creates an antagonist with overgeneralizations and hyperboles: schools are ineffective because "the student must wait perhaps till another day" to know if they are correct, children are "uncertain" and "anxious," and the work is not "pleasurable." Like Pressey, Skinner's answer to those intractable problems is a simple mechanical machine that asks questions and shows correct answers. Again, there is no apparent connection between the stated educational problem (including uncertainty, lack of motivation, and anxiety,) and the proposed technological solution (a question-and-answer machine).

Where Pressey's discourse lauds the mechanization of education—an argument at least in line with the device he designed—Skinner's discourse is instead one of *humanization*: "Most students feel that machine study has compensating advantages. They work for an hour with little effort, and they report that they learn more in less time and with less effort than in conventional ways" (Skinner, as cited in Watters, 2015a).

Less effort, less toil, more learning. Note that even within this humanistic promise, our semiotic toolbox reveals that even Skinner surreptitiously admits his "machine study" is problematic: it has "compensating advantages" rather than advantages unqualified. Perhaps this linguistic slip is a preemptive reaction to the fact that replacing teachers with technology indeed sounds dehumanizing. Accordingly, Skinner is quick to assign a *new meaning* to humanization: "Another important advantage is that the student is free to move at his own pace. When a whole class is forced to move forward together, the bright student wastes time waiting for others to catch up, and the slow student (who may not be inferior in any other respect) is forced to go too fast. . . . he gets farther and farther behind and often gives up altogether" (Skinner, 1954).

In Skinner's discourse, humanization is about less effort, less repetitive work, and, above all, moving at your own speed. But it is never about deviating from the preset curriculum his machine offers. This crucial and consequential discursive move—redefining "humanizing education"—would transform our perception of automated educational technologies. Skinner inverts the normal state of affairs: a machine that forces children to learn standardized content is humanized and becomes an instrument of freedom, while the teacher, portrayed as more mechanical than a machine, becomes an instrument of oppression. His testing machine was no longer about testing but instead about learning:

> [The student] is not in any sense being tested, instead, helpful hints, suggestions, and prompts maximize the chances that he will be right. . . . Programs have been constructed in which without any prior study, the average student is right 95% of the time. This result is partly due to the fact that the student only moves on when he has completely mastered all the preceding material. A conservative

> estimate [is that a] high school student can cover about twice as much material with the same amount of time and effort as with traditional classroom techniques. (Skinner, 1954)

In 1954, Skinner's words were connected to three other intertextual discourses. The first was Pressey's ideas of technical mechanization to overcome the "clumsy procedures" of education. Technology and industry were transforming the world, and the freedom derived from automation (machines freeing people from repetitive tasks) was on everyone's mind. Second, an updated conception of schooling was slowly emerging, one that moved away from Pressey's focus on cognition ("clear thinking") and was instead about motivation, enjoyment, and student-centeredness. A "twisted" version of equity even enters the stage: the "slow student," who would often be deemed genetically inferior and undeserving of any attention during Cubberley's (1909) racist era, now, during Skinnerian times, "may not be inferior in any other respect." Third, by the 1950s, a common notion was that all aspects of human life could be measured and optimized: an increased need for education to be scientifically studied and precisely engineered. The teaching machine, in conversation with these texts, found its place in the minds of many educational reformers: it offered the opportunity to not only mechanize but also measure, rank, and quantify in minute detail.[1] Skinner's machine was not a passing fad—it persisted for decades with multiple versions, and commercial products were sold to the public for years. "Moving at one's own pace," personalization, and individualization would become the cornerstones of automated educational technologies for the next seven decades.

Silicon Valley and the Invention of the Rewind Button

Pressey's and Skinner's discursive moves laid the foundations for the entire industry of automated instruction by formulating justifications and ideas that had, at the same time, the appeal of humanization and automation. This remarkable combination allowed for the apparent resolution of one of the uncomfortable facts of public education—humanization and automation are often at odds with each other due to issues such as cost, teacher-student ratio, and assessment automaticity. When you automate education, most likely you dehumanize it. Because Skinner and Pressey managed to insert a magic device in that equation to accomplish the opposite, their ideas showed remarkable resiliency, and the texts they left behind are periodically revived.

However, it was the Silicon Valley educational awakening of the early 2010s that most effectively brought the teaching machine back into full swing. One of the main actors was Khan Academy, the online video lecture website famously introduced to the world in a 2011 TED Talk. Following the playbook of earlier educational technologists, founder and CEO Salman Khan began by creating the usual antagonist: traditional education.

> A teacher, no matter how good, has to give this one-size-fits-all lecture to 30 students—blank faces. . . . Good students start failing algebra all of a sudden, and start failing calculus all of a sudden, despite being smart, despite having good teachers, and it's usually because they have these Swiss cheese gaps that kept building throughout their foundation. . . . When those teachers are doing that [using the videos] there's the obvious benefit—the benefit that now their students can enjoy the videos in the way that my cousins did, they can pause, repeat at their own pace, at their own time. But the more interesting thing—and this is the unintuitive thing when you talk about technology in the classroom—by removing the one-size-fits-all lecture from the classroom, and letting students have a self-paced lecture at home, then when you go to the classroom, letting them do work, having the teacher walk around, having the peers actually be able to interact with each other, these teachers have used technology to *humanize the classroom*. (emphasis added, Khan, 2011)

By 2011, Skinner had long fallen out of favor—so it is not a coincidence that Khan does not mention him. In the twenty-first century, Skinner's behaviorism represents old, traditional, oppressive schooling: not the type of ideas venture capitalists and philanthropists want to associate with. However, the magic of intertextuality was still doing its job: Khan follows Skinner's ideas with astonishing fidelity—almost word for word. Khan's central claim is that students can rewind video lectures and play them again—not exactly a revolution since the same can be done with a textbook and a plethora of other educational materials. In addition, a student in a teacher-led classroom (face-to-face or remote) can also ask the instructor to repeat a piece of information, and the teacher can always check for understanding. Good teachers know that when students express difficulties understanding the content, they can rephrase the explanation, mention examples, or employ new pedagogical moves. Rarely would a teacher "press the rewind button" and repeat the same sentences, and seldom would a simple repetition word for word lead to deeper understanding (except, possibly, for procedural knowledge). However, Khan's portrait of the classroom downplays all the possibilities available to teachers—or assumes that most teachers do not do it. In his creation of an overgeneralized, stereotyped antagonist, he finds a way to justify his technology.

Nevertheless, the discursive moves here require a more nuanced analysis. We want to avoid oversimplifying Khan as an ill-intentioned designer who is merely copying Skinner's ideas while hiding his inspirations. Like many of the innumerable digital educational repositories in existence, Khan's library of videos can be helpful for students. Some might benefit from these repositories at home when reviewing the day's materials or doing homework—in the same way as they would use a textbook. But this is hardly how the technology of the "rewind button" is portrayed; the narrative needs to adhere to the hyperbolism and transcendentalism of education disruption. Khan was introduced at the TED conference in 2011 by none other than Bill Gates—the stakes were high, and the humble invention of a mere "library of supplemental videos" would not cut it. Polyphony compensated for the invention's simplicity: a TED event, Bill Gates, an MIT-educated "inventor" . . . even before Khan uttered a word, these nonverbal discourses were already elevating the message and setting the stage for success (Benveniste, 1966). As he spoke, Khan picked up texts left "in the air" for decades by Pressey and Skinner. First, he caricaturizes and simplifies what schools are and what teachers do. Then, he supplants the actual classroom—in which (despite all possible and fair criticism) there are humans with multiple possibilities of interaction—replacing it with a video lecture library with a rewind button that always plays the *same* video. Persuasion here is carried out by portraying reality upside down: the human becomes the oppressor, and the machine becomes the humanizer.

Another one of Khan's hallmark ideas is "mastery learning"—again, an idea imported directly from Skinner. Khan says, "Learn math the way you'd learn anything, like riding a bicycle. Stay on that bicycle. Fall off that bicycle. Do it as long as necessary, until you have mastery. The traditional model, it penalizes you for experimentation and failure, but it does not expect mastery. We encourage you to experiment. We encourage you to fail. But we do expect mastery" (Khan, 2011).

The "detective" work of semiotics enables us to unpack his metaphor: learning to ride a bicycle is a vivid example of a skill that is learned in a purely automatic way—in learning it, through trial and error, we develop *no* mechanistic understanding of the science behind it. Learning how to ride a bicycle bears no resemblance to learning how to think mathematically or make sense of mathematical relationships. Here, Khan is betrayed by his own mental model of learning, clearly inspired by the behaviorists. *Experimentation*—a word associated with progressive pedagogies—finds itself here walled in by his system: it becomes merely playing back videos and trying out different answers in quizzes. It removes the fundamental agency of experimentation, with which students can create hypotheses, design experiments to test them, or generate inventions: it gets reduced to being right or wrong until you are "always right." (Remember Skinner [1954]: "Programs have been constructed in which . . . the average student is right 95% of the time. . . . [T]he student only moves on when he has completely mastered all the preceding material.")

Knewton, another Silicon Valley edtech titan, and its former CEO Jose Ferreira made similar claims. Since it came to market a few years after Khan Academy, Knewton's almost identical discourse already used some of the vocabulary of artificial intelligence in education that came into vogue in the late 2010s: "If a student struggles to complete an assignment, our adaptive technology diagnoses and remediates that student's knowledge gaps with personalized content and assessment that will help them achieve proficiency" (Get Results with Alta, 2020).

Intertextuality, again, proves to be a pivotal lens for understanding the discourses of Khan, Skinner, Jose Ferreira, and others. We see the recurrence of terms like "gaps in knowledge," "Swiss cheese gaps," "mastery" and "proficiency," "at your own pace," and "personalization"—even almost 70 years later.

Justin Reich addresses this phenomenon with an illuminating example. While Jose Ferreira of Knewton was claiming that his system was "like a robot tutor in the sky that can semi-read your mind and figure out what your strengths and weaknesses are, down to the percentile," Reich (2020) says, "Knewton engineers were simultaneously publishing blog posts with titles like 'Understanding Student Performance with Item Response Theory.' Lift up the hood of the magical robot tutor, and underneath was a 40-year-old technology powering the whole operation." In other words, Reich reveals that the existing technology used by Knewton (and many other systems) was nothing more than decades-old statistical methods, portrayed as a "robot tutor" in the era of hyperbolic educational discourse.

The Decade of the MOOCs

The pattern of hyperbolic, transcendental discourse elevating trivial innovations was front and center in one of the most spectacular edtech stories of the 2010s: the Massive Online Open Course (MOOC). Daphne Koller, one of the proponents of MOOCs, explained how they were different from "old-fashioned" online learning: "What made these courses so different? After all, online course content has been available for a while. This [a MOOC] was a real course experience. It started on a given day and then students would watch videos on a weekly basis and do homework assignments. And these would be *real* homework assignments for a *real* grade. With a *real* deadline" (emphasis added, Koller, 2012).

We see the same phenomenon repeated yet again: an old technology "powering the whole operation," while the creation of a stereotyped antagonist compensates for the product's simplicity. In touting "classic" online learning as the antagonist for MOOCs, Koller notes that these earlier courses had no start and end dates, homework, or grades, which is not true: much of pre-MOOC online learning was exceedingly formalized. Koller goes on to note that, freed from the constraints of the one-hour classroom lecture, online materials could be broken up into discrete chunks of less than 15 minutes each, leading to customizable options like extra enrichment material. And she continues: "This format allows us to break away from the one-size-fits-all model of education and allows students to follow a much more personalized curriculum" (2012).

After touting MOOCs as a "real" course with "real deadlines," she then characterizes them as the opposite, "break[ing] away from the one-size-fits-all model of education." If the antagonist in the first excerpt were online courses that had no structure, in the second, the antagonist is now courses with structure that follow the "one-size-fits-all model" (which, supposedly, has start and end dates, assignments, and deadlines). Against those, MOOCs are touted as "a much more personalized curriculum." Paradoxically, the product is better at first because it has "real deadlines," then—just a few seconds later—it is good because it is not "one-size-fits-all."

These examples show a pattern in the "folds of the discourse" of automated educational technologies. With most products being based on long-existing technologies (Reich, 2020; Watters, 2021), and with little conceptual or technological innovation, the discourse around them becomes a patchwork of comparisons with an imaginary, stereotyped antagonist that looks conspicuously like the very innovations the new products are trying to replace.

AI: The New Frontier of Mechanized Education

Further evidence of this pattern's entrenchment can be observed in contemporary automated instructionist systems using artificial intelligence, which assume that the abundance of data will shed light onto complex human learning. However, this assumption has often incentivized companies to seek data sources that are easy to manage, regardless of the quality of information they carry about learning or the biases they may contain. Clickstreams and data from online environments provide plentiful, low-cost data, but they render impoverished portraits of how students learn; as Reich (2020) says: "Our assessment technologies are particularly good at assessing the kinds of human performance that we no longer need humans to perform."

Second, in complex social environments such as schools, the number of dimensions of the problem space is so large that the sheer availability of data, without an initial educational theory, will likely never render usable findings. Third, AI in education has a scaling problem: solving an algebra problem is orders of magnitude simpler than designing and executing a scientific experiment or writing a sophisticated essay on a historical topic. Early successes with algebra tutors or arithmetic apps do not justify applying the same student models to more complex topics (Berger, as cited in Hess, 2018).

Many of those assumptions and limitations were behind failed educational start-ups that promoted AI uses for automated instructionist technologies, such as the School of One, AltSchool, and Knewton (Reich, 2020). For those companies, many of the components of good classroom practice, such as social interaction, culturally relevant pedagogy, and pedagogical flexibility, are obstacles for the technologies to work because their algorithms cannot deal with them. These components are then systematically excluded from the systems, and recent studies have shown the ineffectiveness of automating them (see, e.g., a large-scale study of low-cost behavioral manipulations, Kizilcec et al., 2020).

Despite these limitations, the discourse of AI in education marches on. Squirrel Ai Learning, one of the largest AIEd companies globally, claims to have a system with a

> "simulated human teacher giving the student a personalized learning plan and one-on-one tutoring, *with 5 to 10 times higher efficiency*. There is a growing ability to customize the teaching, then students can learn the same amount of material much faster. . . . [It] keeps the students more engaged in a lesson to learn more material in a smaller amount of time. . . . Recognizing from facial expression when the student is happy or bored or frustrated." (emphasis added, Squirrel Ai Learning, n.d.)

The same themes from Skinner and Pressey emerge: *acceleration*, *automaticity*, *efficiency*, and *personalization*, as we will discuss in the next section.

Discussion

Having described the various discursive threads of automated instructionist technologies and found common themes in the data, we comment on each thread and point out how those technologies and their politics impact school systems.

Intertextuality in Action: “Moving at Your Own Pace” and “Mastery Learning”

The several excerpts selected for this chapter show consistency among discourses spanning an entire century. Two are particularly remarkable. The first, in figure 11.1, shows intertextuality in action for the idea of inefficiency and massification of existing educational systems and the “vilification” of lectures. It also shows an astonishing, almost literal permanence of the idea of “moving at one’s own pace,” from Skinner to Khan.

The second set of discourses, in figure 11.2, shows the permanence of two core ideas in behaviorist systems: mastery learning and immediate feedback. The former suggests that students should not progress until they have “mastered” a given topic, as measured

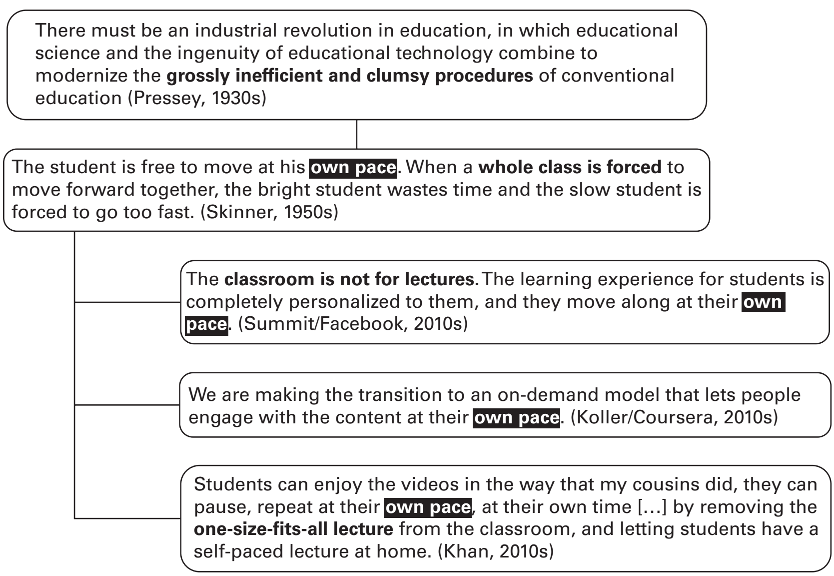


Figure 11.1

Intertextuality spanning one century: from Pressey’s technology to Khan and Summit/Facebook Schools. Note, in the text with the black background, the theme of “learning at your own pace” repeated over seven decades, and, in bold, the creation of the antagonist of edtech: traditional lectures.

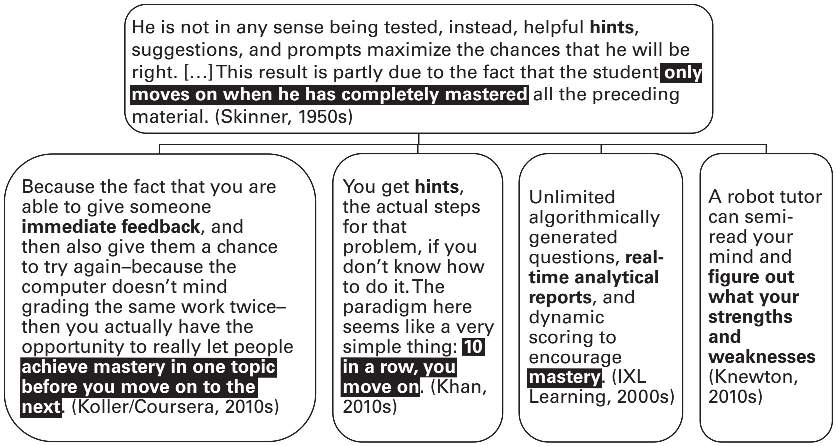


Figure 11.2
Intertextuality in Skinner, Koller, Khan, and AI companies. Note, in bold text, the Skinnerian idea of not tests but "hints" and immediate feedback, and, in the text with the black background, the mastery learning concept: students are only allowed to move on when one topic is "mastered" (IXL, 2020).

by their answers in tests. The latter holds that by repeatedly being tested and receiving rapid feedback, students will learn more and in a "pleasurable" way—equating enjoyable learning with curricular compliance and claiming that answering simple questions is the only valid demonstration of learning.

Both examples illustrate how Pressey's and Skinner's old discourses "traveled in time" almost intact and are still at the foundation of the entire industry of automated instruction (for extensive documentation, see Watters, 2021). Despite their inconsistencies and the lack of evidence-based support, these concepts are not only still efficient in occupying headlines, raising money, and exciting policy-makers; they also show extraordinary resilience.

Neutral, Apolitical Reform: Antagonizing the Stereotyped "Lecture" While Keeping Everything the Same

A clear pattern in the data we mentioned was to stereotype the "lecture" as the root of all problems in education, separating it from the rest of the educational system. Creating this convenient antagonist was much easier than battling the entire system with its obsession for tests, its outdated and overcrowded curriculum, and its oppressive incentive and promotion systems. Fighting such a system would entail critiques of power, privilege, control, racism, and inequality. It is more convenient to limit the problem to a less controversial target: lecturing. In doing so, these technologists selectively appropriate the critique of traditional schooling by progressive educators (Blikstein & Zuffo, 2004), minus the overarching analysis of the historical, economic, and political reasons that generated the "lecture." The task at hand thus became to replace the lecture while keeping the rest of the superstructure intact. No venture capitalist wants to finance the disruption of the social order, complicated discussions about educational and social justice, debates about schooling's social function, or costly rewritings of national standards. However, since the born-again online lecture has to be different, technologists re-signify frivolous design elements as revolutions, adding "novelties" such as the ability to rewind videos or gamification. To elevate such trivialities to the level of genuine educational innovations, polyphony is recruited: famous personalities are brought on board, TED Talks are delivered, and the specialized press is mobilized. Yet the enabling social systems that gave rise to the vilified lecture remain intact and unquestioned, all but ensuring that today's edtech products will become little more than scapegoats to justify funding tomorrow's "innovations."

Acceleration of Learning and the Lost Einsteins

Beyond its dishonesty, the discourse of automatic instructionist technology reveals assumptions that contradict the principles of equity and inclusion purportedly at the heart of modern educational systems. A recurrent topic in the data is the *acceleration* of learning. Skinner, Khan, Ferreira, and Koller mention "learning more in less time" as a core goal. Focusing on *productivity* and *speed* avoids conversations about giving students more agency, making curricular topics more flexible, or attending to culturally meaningful standards. “Learning at your own pace” is never associated with learning about topics of personal interest—for all of Khan’s or Skinner’s concern with equity, the freedom given to the “slow student” is simply to catch up with the class by devoting more hours at home to study the same topic. The “personalization” refers only to the video playback speed or the ability to pause and rewind—not the agentic type of personalization of John Dewey or Paulo Freire that generates deep learning and engagement.

The “acceleration” of learning is, in fact, a cyberspace version of school tracking, a practice in which students are sorted into different classrooms and life trajectories by performance. The rhetoric is revealing: Khan (2011) talks about slow students being allowed to “repeat videos at their own pace, at their own time,” and Skinner (1954) about the “bright student wast[ing] time waiting for others to catch up.” What is expected from the “slow ones” is mere curricular compliance, done in their own time, not “our” time (in other words, the “slow” students should remediate their learning speed on their own). Indeed, in many systems worldwide, school systems—especially in low-resourced areas—are increasingly (and unfortunately) relying on automated electronic resources for remedial education.

The other side of this focus on tracking is revealed by how often entrepreneurs talk about how we are “losing Einsteins” by offering low-quality education around the world (Khan, 2011). Notwithstanding the poor choice of Einstein as an example (he famously begrudged schooling), the political stance comes full circle: while we send the “slow” students home to catch up on the day’s lesson on their own time, we look forward to the “Einstein-level” students who might bring real value to society through their genius.

The categorization of “slow” and “fast” students again echoes Pressey’s obsession with intelligence tests and social sorting, and it reveals an educational mindset that opposes ideas of equity and inclusion. Again, the similarity of the "lost genius" discourse, the obsession with sorting students into speed-of-learning categories, and technology being able to accelerate learning are another set of well-aligned discourses that connect, through intertextuality, Pressey and Skinner with their modern instantiations.

Changing the Nature of Learning: Educational Soylent

The significant dissonance between the publicized goals and the real solutions of edtech's automated instruction has been apparent since Pressey's claim that his mechanical contraption for testing could generate "clear thinking, and high ideals." Even Larry Berger, the CEO of Amplify, one of the largest companies in automated instruction, recognized in 2018 that:

> I was a great believer in "personalized learning." . . . Here's the problem: The [learning] map doesn't exist, the measurement is impossible, and we have, collectively, built only 5% of the library. . . . [It] doesn't exist for reading comprehension, or writing, or for the more complex areas of mathematical reasoning, or for any area of science or social studies. The existing measures are not high enough resolution to detect the thing that a kid should learn tomorrow. (Berger as cited in Hess, 2018)

Berger's recognition reinforces the hypothesis that edtech's only salvation is not to improve education as we know it but to *change its very nature*, transforming it into something easily automated. However, education, at least in its current form, is anything but automatable. As an analogy, take a family's rituals for preparing or eating meals. They serve multiple educational, social, and psychological purposes. It would be challenging to create products to replace all of them unless you re-signify eating and make it about eliminating the feeling of hunger. That is precisely what a well-known start-up company, Soylent, did in the mid-2010s. It created an "all-in-one" shake containing 33 percent of a human's nutrient needs and advertised that three bottles were all we needed to survive. To be successful, Soylent had to make customers believe that cooking and eating were a waste of time and that food was just about efficiently consuming essential nutrients. Its then CEO Rob Rhinehart famously said that supermarkets were "endless confusing aisles [with] the smell of rotting flesh," and kitchens were akin to torture chambers with "red hot heating elements and razor-sharp knives." In other words, the company had to transform cooking and eating into an inefficient, inconvenient, and dangerous experience. The company eventually changed its mission and ousted the CEO but not before raising 75 million dollars and occupying headlines for two years (McAlone, 2015).

Consequently, the entire project of automated instruction in edtech can succeed only if the nature and purpose of education itself changes. Such a change requires erasing the socialization of children in schools, noncurricular learnings, unquantifiable knowledge, complex facilitation, play, inquiry, curiosity, public control, and other aspects of education that we cannot measure, package, and automate.[2] By narrowing our focus to straightforward content and multiple-choice tests (even those disguised as "personalized"), automation becomes a much easier task. It depends on undervaluing and ultimately rendering invisible the work that teachers do in classrooms beyond content recitation—because that work is precisely what automated instructionist technologies cannot do.

Furthermore, overlooking this nonautomatable work could further exacerbate existing educational inequities. AI-based or automated educational systems, with their dependence on vast amounts of easy-to-discretize data, are more cost effective for specific topics within STEM disciplines. Teaching for complex problem-solving, exploring multidimensional phenomena, or learning outside STEM will increasingly be outside the realm of these technologies. It could very well be that resource-strapped public schools adapt to what low-cost automated instruction can do, feeding their students glorified educational Soylent, while affluent schools continue to offer rich, complex, hard-to-automate learning experiences to their students. Due to its intrinsic technological limitations, automated teaching could be a harmful tool that would not only communicate but also ossify children's place in society and deny them access to symbolic capital that is already sparse in curricular content (Anyon, 1980)—and even less so when the curricula are automated.

Conclusion: The Politics and Enabling Systems of Automated Teaching Technologies

Our analysis revealed a familiar pattern in the discourses of teaching technologies. Namely, entrepreneurs propose a new automated educational technology by establishing an opposition to a stereotyped version of traditional education (*dialogism*). Then, they build on *intertextuality* to generate discourse that makes use of old and new "texts" (e.g., "learning at your own pace"). Finally, through *polyphony* (social media, marketing, high-profile events, celebrity endorsements, branding), they disseminate and legitimize the inevitability of the product—which is then assimilated into everyday discourse (e.g., "personalized learning" is now incorporated into the lexicon of schools and policy-makers). This assimilation is often unproblematic because the politics of these technological artifacts (which advance an agenda of standardization, ranking, tracking, optimization, and efficiency) operate in a matching enabling system that, despite appearances, reinforces and welcomes this agenda. Ultimately, automated technologies can help satisfy the need for novelty without changing the systems they're used in.

The sequence of events is strikingly consistent: identify an antagonist ("traditional education") through dialogic discourse, propose a technical solution unrelated to real educational issues, frame your technology as “more humanist than humans,” overstate the innovation, and obfuscate the solution’s simplicity. Through this process, proponents of automated instructionist technologies do not expand their products’ use through traditional means—that is, pilot projects followed by evidence-based research—leading to increasingly larger implementations. Instead, they aim to reclaim the role of innovators in education and, in that capacity, to displace other education stakeholders from setting the agenda, to silence dissenting voices, and to reshape education in the image of the simplistic technologies they produce.

Among the benefits to technologists is the privilege of avoiding critique if their educational formulations go wrong. Instead, edtech entrepreneurs who fail catastrophically are allowed to “pivot” to a different direction without consequence. Take, for example, Coursera’s several “reinventions,” the failure of Udacity’s MOOCs, the ruin of the School of One, Summit Learning, AltSchool, Edmodo, inBloom, and Knewton. And despite being behind many of those failed initiatives, and protests from teachers’ union leaders, in May 2020, the Gates Foundation was announced as the state of New York’s leading partner in “reimagining” education during the pandemic (Blad, 2020).

The semiotic instance here is crucial: by controlling the narrative in this way, failures can be rebranded as humble “pivots,” allowing foundations and entrepreneurs to maintain their statuses as successful innovators (for examples, see Wan, 2016, which recounts interviews of departing or current edtech CEOs). In Benveniste’s terms, the goal of such discourse is not to inform or describe a product’s functionality but instead to produce a positive reaction in the mind of its “customers.” To quote a longtime industry analyst: “Companies like Knewton and others went straight into black-box algorithms. Their customers were really venture capitalists, not academic programs with real teachers and students” (Hill, as cited in Ubell, 2019).

In shining the light of discourse analysis on the world of automated educational technologies, we gain new insight into how these leaders and companies managed to dominate the discourse of educational technologies for decades. We understand why philanthropic foundations and governments invest in products that do not even exist, trust entrepreneurs with no educational experience, and commit national education systems to fragile projects without consistent plans.

Pressey, Skinner, Khan, and Ferreira all contributed to a 100-year-old project to mold education in the image of their technologies. The modern edtech version of behaviorism understands that the actual battle in education centers on a narrative of innovation, disruption, and revolution. With each new technology (e.g., AI), its claims get increasingly hyperbolic at the same time as it becomes increasingly furtive about its theoretical inspirations.

Especially during the COVID-19 pandemic, we realized that these edtech mirages are deeply consequential for children. In a time of increasing social inequality and escalating tensions due to multiculturalism and immigration, automated and AI-based educational systems—in their current inception—could become the ultimate tool for educational stratification and inequity. Such systems could be the tool of choice for low-income and underprivileged school districts due to constant budget pressures and the allure of a Silicon Valley–esque revolution. Students in those districts would not only be exposed to less face-to-face, innovative instruction but would also be much more vulnerable to bias and to having their data exploited or monetized by service providers. These populations would grow up with dehumanizing, impersonal educational technologies that would greatly diminish their prospects in the complex and interconnected world of the twenty-first century.

But it would be a mistake to simply ascribe ill intention to edtech companies and comfortably sit in our academic offices dispensing criticism. We have another job to do. It is up to us to build defenses in our educational systems that will guard against the seductive discourses of automated instructionist technologies. Part of this work lies in ensuring that our educational systems take advantage of technology in other ways instead—such as engaging children in building inventions, programming computers, composing music, questioning the social order, or creating art. Those uses of technology are directly opposed to automation—they need experienced teachers, time, and effort. Rather than focusing on accelerating learning, they make possible the learning of new, previously unthinkable topics and skills. And they are the types of activities that genuinely embody "personalization," enabling children to express their ideas and intellectual passions. This is the *true personalization* that proponents of automated instruction stole. Our role now is to reclaim it.

Acknowledgments

The authors would like to thank Yipu Zheng for her invaluable data collection work, as well as Alicja Żenczykowska, Leah Rosenbaum, Fabio Campos, and Chelsea Villareal, for their insightful comments and reviews of earlier drafts.

Notes

1. While certainly not the first to create such learning machines, Skinner promised that with his machine, "a high school student can cover about twice as much material with the same amount of time." In his defense, at times, he was transparent about what the machine was about—a glorified textbook. In a buried comment toward the end of the film presentation, he comments: "There is no magic about this teaching machine . . . it is simply a convenient way of bringing the student into contact with the man who writes the program" (Skinner, 1954).

2. For examples of how overmeasurement in medicine and business can decrease the quality of services, see Muller (2018).

References

Anyon, J. (1980). Social class and the hidden curriculum of work. *Journal of Education*, *162*(1), 67–92. https://doi.org/10.1177/002205748016200106

Bakhtin, M. (1984). *Problems of Dostoevsky's poetics*. University of Minnesota Press.

Bakhtin, M., Holquist, M., & Liapunov, V. (1993). *Toward a philosophy of the act*. University of Texas Press.

Benveniste, E. (1966). *Problèmes de linguistique générale*. Gallimard.

Blad, E. (2020, December 3). New York state teams with Gates Foundation to "reimagine education" amid pandemic. *Education Week*. https://www.edweek.org/education/new-york-state-teams-with-gates-foundation-to-reimagine-education-amid-pandemic/2020/05

Blikstein, I. (2020). *Semiótica e totalitarismo*. Contexto.

Blikstein, P., & Zuffo, M. K. (2004). As sereias do ensino eletrônico [The mermaids of electronic teaching]. In M. Silva (Ed.), *Educação Online: Teoria, prática, legislação e treinamento corporativo [Online education: Theory, practice, legislation and corporate training]* (pp. 10–37). Ed. Loyola.

Buechley, L., & Eisenberg, M. (2009). Fabric PCBs, electronic sequins, and socket buttons: Techniques for e-textile craft. *Personal and Ubiquitous Computing*, *13*(2), 133–150.

Cole, R., Kemple, J. J., & Segeritz, M. D. (2012). *Assessing the early impact of School of One: Evidence from three school-wide pilots*. Research Alliance for New York City Schools, New York University. https://steinhardt.nyu.edu/research-alliance/research/publications/assessing-early-impact-school-one

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. Teachers College Press.

Cubberley, E. P. (1909). *Changing conceptions of education*. Houghton Mifflin.

DiSalvo, C. (2014). Critical making as materializing the politics of design. *The Information Society*, *30*(2), 96–105. https://doi.org/10.1080/01972243.2014.875770

Freire, P. (2014). *Educação como prática da liberdade*. Editora Paz e Terra.

Get Results with Alta. (2020). Knewton. https://www.knewton.com/results/

Ginzburg, C. (1991). Chaves do mistério: Morelli, Freud e Sherlock Holmes. In U. Eco & T. A. Sebeok (Eds.), *O signo de três* (pp. 89–129). Perspectiva.

Hess, R. (2018, February 12). A confession and a question on personalized learning [Opinion]. *Education Week*. https://www.edweek.org/leadership/opinion-a-confession-and-a-question-on-personalized-learning/2018/02

IXL. (2020, August 4). *IXL Analytics*. https://www.ixl.com/analytics

Jakobson, R. (1968). *Lingüística e comunicação*. Cultrix.

Joerges, B. (1999). Do politics have artefacts? *Social Studies of Science*, *29*(3), 411–431.

Khan, S. (2011, March 9). *Let's use video to reinvent education* [Video]. TED Conferences. https://www.ted.com/talks/sal_khan_let_s_use_video_to_reinvent_education?language=en

Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., Turkay, S., Williams, J. J., & Tingley, D. (2020). Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences*, *117*(26), 14900–14905. https://doi.org/10.1073/pnas.1921417117

Koller, D. (2012, August 1). *What we're learning from online education* [Video]. TED Global. https://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education

Latour, B. (2004). Which politics for which artifacts. *Domus, 171(June)*, *4*, 50–51.

McAlone, N. (2015, October 15). *24 controversial quotes from Soylent's CEO that will either terrify or inspire you about the future of food*. Business Insider Australia. https://www.businessinsider.com.au/soylent-ceo-rob-rhinehart-quotes-about-the-future-of-food-2015-10

Moles, A. A. (1958). *Théorie de l'information et perception esthétique*. Edition Flammarion.

Muller, J. Z. (2018). *The tyranny of metrics*. Princeton University Press.

Papert, S. A. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Ready, D., Conn, K., Bretas, S., & Daruwala, I. (2019). *Final impact results from the i3 implementation of Teach to One: Math*. Consortium for Policy Research in Education (CPRE) at Teachers College, Columbia University. https://www.tc.columbia.edu/media/centers/cpre/Final-Impact-Results-i3-TtO-190224.pdf

Reich, J. (2020). Two stances, three genres, and four intractable dilemmas for the future of learning at scale. In *Proceedings of the Seventh ACM Conference on Learning @ Scale* (pp. 3–13). ACM. https://doi.org/10.1145/3386527.3405929

Selwyn, N. (2010). Looking beyond learning: Notes towards the critical study of educational technology. *Journal of Computer Assisted Learning*, *26*(1), 65–73. https://doi.org/10.1111/j.1365-2729.2009.00338.x

Selwyn, N. (2013, May 23–25). *Discourses of digital 'disruption' in education: A critical analysis*. Paper presented to Fifth International Roundtable on Discourse Analysis, City University, Hong Kong. https://www.academia.edu/4147878/Discourses_of_digital_disruption_in_education_a_critical_analysis

Selwyn, N., & Facer, K. (2013). Recognizing the politics of 'learning' and technology. In N. Selwyn & K. Facer (Eds.), *The politics of education and technology: Conflicts, controversies, and connections* (1st ed.). Palgrave Macmillan.

Skinner, B. F. (1954). *Teaching machine and programmed learning* [Video]. YouTube. https://www.youtube.com/watch?v=jTH3ob1IRFo

Squirrel Ai Learning. (n.d.). *Our story*. Retrieved September 10, 2021, from http://squirrelai.com/our-story

Todorov, T. (1989). *Nous et les autres. La réflexion française sur la diversité humaine*. Éditions du Seuil.

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Harvard University Press.

Ubell, R. (2019, June 12). *Explaining the shakeout in the adaptive learning market (opinion)*. Inside Higher Ed. https://www.insidehighered.com/digital-learning/views/2019/06/12/explaining-shakeout-adaptive-learning-market-opinion

Vakil, S. (2018). Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review*, *88*(1), 26–52.

Wan, T. (2016, December 21). *Jose Ferreira steps down as Knewton CEO, eyes next education startup*. EdSurge. https://www.edsurge.com/news/2016-12-21-jose-ferreira-steps-down-as-knewton-ceo-eyes-next-education-startup

Watters, A. (2015a, February 10). *Education technology and Skinner's box*. Hack Education. http://hackeducation.com/2015/02/10/skinners-box

Watters, A. (2015b, February 19). *The history of the future of education*. Hack Education. http://hackeducation.com/2015/02/19/the-history-of-the-future-of-education

Watters, A. (2021). *Teaching machines: The history of personalized learning*. The MIT Press.

Williamson, B. (2018a). Silicon startup schools: Technocracy, algorithmic imaginaries and venture philanthropy in corporate education reform. *Critical Studies in Education*, *59*(2), 218–236. https://doi.org/10.1080/17508487.2016.1186710

Williamson, B. (2018b, March 29). *Why education is embracing Facebook-style personal profiling for schoolchildren*. The Conversation. https://theconversation.com/why-education-is-embracing-facebook-style-personality-profiling-for-schoolchildren-94125

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121–136. http://www.jstor.org/stable/20024652

Winner, L. (2009). Information technology and educational amnesia. *Policy Futures in Education*, *7*(6), 587–591. https://doi.org/10.2304/pfie.2009.7.6.587

Woolgar, S., & Cooper, G. (1999). Do artefacts have ambivalence: Moses' bridges, winner's bridges and other urban legends in S&TS. *Social Studies of Science*, *29*(3), 433–449.

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*(1), 75–89. https://doi.org/10.1057/jit.2015.5

Contributors

Izidoro Blikstein was a professor at the Department of Romance Linguistics at University of São Paulo, Brazil. He is the author of *Kaspar Hauser or the Fabrication of Reality*, and his research focuses on the semiotics of authoritarian regimes.

Paulo Blikstein is an associate professor at Teachers College and (by affiliation) at the Computer Science Department at Columbia University, EUA. His research and design focus on emancipatory technologies for education. https://tltlab.org

Marion Boulicault is a distinguished postdoctoral fellow in ethics and technology at the MIT Schwarzman College of Computing, a founder of the Experiential Ethics program at MIT, and the director of interdisciplinary research at the Harvard GenderSci Lab. Her research focuses on feminist philosophy of science and technology.

Cynthia Breazeal is a professor of media arts and sciences at the Media Lab. Her research is about social robots and their responsible use, spanning technical innovations, user-centered design, and the psychology of engagement with focus on education, health, aging, and wellness application domains.

Michelle Ciccone is a PhD student in communication at the University of Massachusetts Amherst. Previously she was a K-12 technology integration specialist, and she continues to build media literacy and critical digital citizenship curriculum for middle and high school grades. Michelle is affiliated with the Media Education Lab at the University of Rhode Island, was a research assistant at the Berkman Klein Center for Internet & Society at Harvard University, and in 2019 was awarded third place in AEJMC's Best Practices in Teaching Media Literacy competition.

Remy Cross is a sociologist and a research manager at the Connected Learning Lab. He was a professor and department chair at a small Midwestern university before joining the Connected Learning Lab to assist in its mission. His previous work has examined violent radicalization, particularly in online environments, as well as work on online communities and collective action. Remy holds a PhD in sociology from the University of California, Irvine as well as master's degrees from the University of Chicago and University of California, Irvine.

Sayamindu Dasgupta is an assistant professor at the Department of Human Centered Design & Engineering, University of Washington (UW). He designs and studies new ways young people can learn with and about data—especially in contexts of the communities that they live, learn, and play in. https://unmad.in/

Devin Dillon, senior director of Technovation, leads AI Family Challenge at Iridescent. In the program's first year, Devin scaled the initiative into 13 countries with 7,500-plus participants.

Karthik Dinakar is a computer scientist in the fields of natural language processing, machine learning, and clinical medicine.

Stefania Druga is the creator of Cognimates, a platform for AI education for families, and a PhD student at the University of Washington. Her research on AI education started during her master's in Personal Robots Group at MIT Media Lab. She is also a former Weizenbaum research fellow in the Critical AI Lab and an assistant professor at New York University's Interactive Telecommunications Program and Rhode Island School of Design, teaching graduate students how to hack smart toys for AI education.

Benjamin Mako Hill is an assistant professor of communication at the University of Washington and a faculty associate at the Berkman Klein Center for Internet and Society at Harvard University. He studies digital public goods, collective action, and collaborative learning in online communities. https://mako.cc/academic/

Mizuko Ito is a cultural anthropologist, a learning scientist, an entrepreneur, and an advocate for connected learning—learning that is equity oriented, centered on youth interest, and socially connected. Her work decodes digital youth culture for parents and educators, offering ways to tap interests and digital media to fuel learning that is engaging, relevant, and socially connected. She is the director of the Connected Learning Lab and professor in residence and John D. and Catherine T. MacArthur Foundation Chair in Digital Media and Learning at the University of California, Irvine. She is also cofounder of Connected Camps, a nonprofit providing online learning experiences for kids in all walks of life. Her most recent publications include *Affinity Online: How Connection and Shared Interest Fuel Learning*, *From Good Intentions to Real Outcomes: Equity by Design in Learning Technologies*, *The Connected Learning Research Network: Reflections on a Decade of Engaged Scholarship*, and *Social Media and Youth Wellbeing*.

Jacqueline M. Kory-Westlund is an independent scholar with the Ronin Institute, a writer, and an artist with a PhD from the MIT Media Lab. Her research has focused on using social robots to support and engage young children in learning, including how children understand social robots, how children's relationships connect to their learning, especially over long-term interactions, and the ethics of using robots in children's lives.

Aviv Y. Landau, PhD, MSW worked for several years as a social worker in Israel and is a postdoctoral research scientist at the School of Social Practice and Policy and codirector of the SAFElab, University of Pennsylvania.

Adriana Manago is an associate professor of developmental psychology at the University of California, Santa Cruz. Research in her Culture and Technology lab examines communication technologies, cultural change, and social and identity development from adolescence through the transition to adulthood in diverse cultural communities, including a Maya community in southern Mexico.

Siva Mathiyazhagan is a postdoctoral research scientist and associate director of Strategies and Impact at SAFElab, University of Pennsylvania. His research focuses on a youth-led transdisciplinary approach to prevent racial bias and misinterpretation of social media posts by BIPOC youth to promote safe and inclusive off- and online communities. He co-develops tech social work policies and social work ethical considerations for artificial intelligence (AI), machine learning (ML), and Metaverse. Mathiyazhagan is the founding director of Trust for Youth and Child Leadership TYCL International, and TYCL representative to the United Nations.

Maureen Mauk is a doctoral candidate in Media and Cultural Studies at University of Wisconsin's Communication department. Mauk studies the intertwined relationship between parents, policy, and industry as it relates to television history and the current platformized media landscape. She carries a decade of experience serving in Los Angeles as a Television Standards & Practices executive and has been published in several journals including *Journal of Cinema and Media Studies*, *Learning, Media & Technology*, and *Spectator*.

Stephanie Nguyen is a human-computer interaction designer and researcher, specializing in data privacy, user experience design, and tech policies that impact vulnerable populations. She is an appointed member of IEEE Standards Association's Global Advisory Council on Children's Experiences and a trustee at 5Rights Foundation focused on children's digital rights and data literacy.

W. Ian O'Byrne is an associate professor of literacy education at the College of Charleston in South Carolina. His research focuses on the dispositions and literacy practices of individuals as they read, write, and communicate in online and hybrid spaces. His work can be found on his website (https://wiobyrne.com/) or in his weekly newsletter (https://digitallyliterate.net/).

Candice Odgers is a professor whose team has been capturing the daily lives and mental health of adolescents using mobile phones and sensors over the past decade. More recently, she has been leveraging digital technologies to better support the needs of children and adolescents as they come of age in an increasingly unequal and digital world. She is the author of over 100 scientific publications, and her research has been disseminated widely via outlets such as the *Economist*, the *New York Times*, *Scientific American*, and the *Washington Post*.

Kathleen A. Paciga is an associate professor of education in the Department of Humanities, History, and Social Sciences at Columbia College Chicago. Katie's work examines the ways in which media, in all its diverse forms, are integrated into children's literate lives as well as how these affect the ways children learn and grown-ups teach. Katie served as the 2020 chair of the Excellence in Early Learning Digital Media Award sponsored by the Association for Library Services for Children.

Desmond U. Patton is the Brian and Randi Schwartz University Professor, with joint appointments in the School of Social Policy & Practice and the Annenberg School for Communication, along with a secondary appointment in the department of psychiatry in the Perelman School of Medicine, University of Pennsylvania. He is the founding director of the SAFElab. Professor Patton is a Public Interest Technologist who uses qualitative and computational data collection methods to examine the relationship between youth and gang violence and social media; how and why violence, grief, and identity are expressed on social media; and the real-world impact these expressions have on wellbeing for low-income youth of color.

Milo Phillips-Brown is an associate professor of philosophy at the University of Oxford Faculty of Philosophy and Department of Computer Science, a tutorial fellow at Jesus College, and a senior fellow in digital ethics and governance at the Jain Family Institute. His research is about the ethics of technology, ethical engineering pedagogy, and the philosophy of mind and language.

Michael Preston (Joan Ganz Cooney Center) is the executive director of the Joan Ganz Cooney Center at Sesame Workshop, a research and innovation lab that works to advance children's learning and health in the digital age. His work has focused on using technology to improve teaching and learning, drive student agency and interest, and create models for systemic change in K-12 and university contexts.

Stephanie M. Reich is a professor in the School of Education with additional appointments in informatics and psychological science, director of the Development in Social Context (DISC) Lab, and core faculty in the Connected Learning Lab at the University of California, Irvine. Her research focuses on understanding and improving the social context of children's lives through exploring and intervening with the direct and indirect influences on the child, specifically through family, digital, and school environments.

Nicholas D. Santer is a doctoral student in developmental psychology at the University of California, Santa Cruz. His current research focuses on adolescents' and emerging adults' lived experiences with communication technologies and social media, focusing on the ways in which youth construct stories of the self across diverse polymedia landscapes.

Allison Starks is a doctoral student in the School of Education at the University of California, Irvine. Her research examines the affordances and limitations of technology for positive child outcomes across contexts, with a special emphasis on school-level technology integration.

Elizabeth Y. Stevens is an associate professor of teacher education at Roberts Wesleyan College. Elizabeth's research interests include literacy teacher education, literacy teacher identity, and literacy and technology. Elizabeth is an area chair for

Literacy Assessment, Evaluation, and Public Policy and an e-editor for the Literacy Research Association.

Kristen Hawley Turner is professor and director of teacher education in the Caspersen School of Graduate Studies at Drew University in New Jersey. She is the coauthor of *Connected Reading: Teaching Adolescent Readers in a Digital World*, coauthor of *Argument in the Real World: Teaching Students to Read and Write Digital Texts*, and the editor of *Ethics of Digital Literacy: Developing Knowledge and Skills across Grade Levels*.

Veena Vasudevan is an assistant professor of digital media and learning at University of Pittsburgh's School of Education. Her broader research agenda explores STEM learning, educational equity, and urban education with an emphasis on understanding the lives of children and youth through the lenses of identity, learning, and literacies. She is currently launching new research into the effects of production-centered critical pedagogies on preservice educators and the role of critical data literacies within K-12 education.

Jason Yip (Digital Youth Lab at the University of Washington) is an associate professor at the Information School and an adjunct assistant professor in the Department of Human Centered Design and Engineering at the University of Washington. His research examines how technologies can support parents and children learning together. His research is supported by the National Science Foundation, the National Institutes of Health, Mozilla, and the Institute of Museum and Libraries.

Index

Note: page numbers in *italics* indicate figures or tables.