1-1-2017

Do Androids Dream of Bad News?

Heidi Liu

Do Androids Dream of Bad News?

Heidi Liu[*]

Abstract

Breaking bad news is one of the toughest things to do in any field dealing with client care. As automation and technology increasingly interweave with human experience, there is growing concern about whether automated agents ("AAs") would be adequate to perform such a complex emotional act. In this paper, I draw from the literature in psychology and computer science to understand how individuals might react to automated agents (AAs) and address some of the strengths and limitations of AAs. I raise several legal and empirical issues that future designers and users of AAs must consider, including disclosure of and liability for an AA's presence.

INTRODUCTION

A recent, widely shared blog post contends that "2016 will be the year of conversational commerce,"[1] where established app companies will allow for direct communication between individuals and businesses.[2] The author, Chris Messina, states that:

> I'm less interested in whether a conversational service is provided by a human, bot, or some combination thereof. If I use these terms interchangeably, it's not *unintentional*. It's just that over an increasing period of time, computer-driven bots will become more human-feeling, to the point where the user can't detect the difference, and will interact with either human agent or computer bot in roughly the same interaction paradigm.[3]

This paper challenges the implication that bot-human and human-human conversations are interchangeable, using one particular interaction: revealing (or "breaking") bad news. Understanding how bad news is received matters because of its downstream implications: a patient might not follow her doctor's orders if she doesn't trust the doctor, or doesn't know how to react to the news of a positive diagnosis. Similarly, a consumer may angrily stop using a health app if it reveals that he is not as healthy as he had believed.

---

[1] Chris Messina, "2016 Will Be the Year of Conversational Commerce," *Medium* (19 January 2016), online: < www.medium.com >.

[2] See *ibid.*

[3] *Ibid.*

The potential of having an AA reveal bad news stems from alleviating the discomfort of both messenger and receiver. If a human messenger overestimates the effect of "shooting the messenger," then she may withhold important information from the recipient. On the other hand, if the human messenger underestimates an effect, she may not only experience backlash but the recipient may also disregard the message based on the messenger's delivery. An AA messenger might be able to break the news in a sensitive way.

While it seems likely that individuals might interact with AAs and humans similarly, are there situations in which this interaction would be inappropriate? And if otherwise indistinguishable, *should* humans know whether an automated agent or human is on the other end? These questions have legal implications, both in terms of how to regulate the revelation of bad news and liability-shifting. In this paper, I examine two examples where automating bad news complicates how we already deliver news in an "offline" context: a lawyer's role in mediation and settlement processes and medical disclosures.

We know little about breaking bad news, as much of the relevant literature takes place in the medical realm, given the high stakes involved. There, the only empirical result suggests that participants want bad news before they receive good news,[4] though to my knowledge, no existing publications look to whether "shooting the messenger" actually occurs.[5] Doctors distribute best practices to inform patients about positive diagnoses, including the provision of "content, context and support" at the time of diagnosis[6] and the likelihood of successful treatment.[7] However, few of these techniques have been empirically validated. Rather, the literature emphasizes how physicians can overcome their reluctance to report bad news[8] — a key motivator for why firms and individuals might want to use an automated messenger. Indeed, although electronic communication appears to provide a shield to facilitate a bad news messenger's honesty and accuracy, little research exists on how people respond to electronic forms of bad news.[9]

---

[4] See Angela M. Legg & Kate Sweeny, "Do You Want the Good News or the Bad News First? The Nature and Consequences of News Order Preferences" (2013) 40:3 Personality & Soc. Psych. Bull. 279 at 285-286.

[5] Leslie John, Hayley Blunden, and I have a series of experiments demonstrating this effect. See Leslie John, Hayley Blunden and Heidi Liu, "Shooting the Messenger" [unpublished].

[6] See Patricia A. Parker et al., "Breaking Bad News About Cancer: Patients' Preferences for Communication" (2001) 19:7 J Clin Oncol 2049 at 2049.

[7] See generally Penelope E. Schofield & Phyllis N. Butow, "Toward Better Communication in Cancer Care: A Framework for Developing Evidence-based Interventions" (2004) 55:1 Patient Educ Couns 32 at 32—39.

[8] See Sidney Rosen & Abraham Tesser, "Fear of Negative Evaluation and the Reluctance to Transmit Bad News" 22:2 J. Comm. 124 at 137; Ahmet Uysal & Bengi Oner-Ozkan, "A Self-Presentational Approach to Transmission of Both Good and Bad News" (2007) 35:1 Soc Behav Pers 63 at 63–64.

[9] But with regard to messengers, people deliver bad news more honestly and accurately

The Computer as Social Agents (CASA) model suggests that humans are likely to interact with AAs in a similar way as humans.[10] Humans are more likely to take advice from agents who appear similar to them, whether in appearance[11] or in personality.[12] Conversely, AA messengers who gave "bad news" to the participant regarding the participant's performance were found to be less agreeable, likable and more dissimilar,[13] although this effect was attenuated when the robot appeared less humanoid.[14] In fact, participants who identify as relational (relative to utilitarian) appear to prefer an apology from a robot rather than compensation.[15]

These negative effects of bad news are amplified when humans work together with agents toward a common task. In another experiment, one team programmed a robot to blame its poor performance on itself or the human participants that had subsequently trained it. Of course, the participants preferred the self-blaming robots.[16] And in fact, in another study, participants blamed the agent for their own performance.[17] Similar reactions occur in even everyday interactions: for instance, when an automated agent informs a participant that a particular timeslot is unavailable, the human's attitude toward the machine does change, albeit slightly.[18]

---

through computer mediated communication compared to phone or face-to-face methods. See Stephanie W. Sussman & Lee Sproull, "Straight Talk: Delivering Bad News Through Electronic Communication" (1999) 10:2 Info Sys Res 150 at 161–162.

[10] See Robert G. Magee & Sriram Kalyanaraman, "The Perceived Moral Qualities of Web Sites: Implications for Persuasion Processes in Human-Computer Interaction" (2010) 12:2 Ethics & Info Tech 109 at 110.

[11] See Ian Li et al., "My Agent as Myself or Another: Effects on Credibility and Listening to Advice" (Paper delivered at the International Conference of Designing Pleasurable Prod & Interfaces, 22-25 August 2007) (2007) DPPI 2007 — Proceedings of the 2007 International Conference of Designing Pleasurable Prod. & Interfaces 194 at 195.

[12] See Adam Waytz, John Cacioppo & Nicholas Epley, "Who Sees Human?: The Stability and Importance of Individual Differences in Anthropomorphism" (2010) 5:3 Persp on Psych Sci 219 at 226—227; Magee & Kalyanaraman, *supra* note 11 at 111.

[13] Leila Takayama, Victoria Groom, & Clifford Nass, "'I'm Sorry, Dave: I'm Afraid I Won't Do That': Social Aspects of Human-Agent Conflict" (Paper delivered at SIGCHI Conference on Human Factors in Computing Systems, 7 April 2009) (2009) Proc SIGCHI Conf on Hum Factors in Computing Syst 2099 at 2105.

[14] *Ibid* at 2106.

[15] Min Kyung Lee et al., "Gracefully Mitigating Breakdowns in Robotic Services" (Paper delivered at ACM/IEEE Int'l Conf, on Hum-Robot Interaction, 4 March 2010) (2010), ACM/IEEE Int'l Conf. on Hum-Robot Interaction 203 at 203.

[16] Victoria Groom et al., "Critic, Compatriot, or Chump?: Responses to Robot Blame Attribution" (Paper delivered at ACM/IEEE Int'l Conf. on Hum-Robot Interaction, 4 March 2010) (2010), ACM/IEEE Int'l Conf. on Hum-Robot Interaction 211 at 215.

[17] See Taemie Kim & Pamela Hinds, "Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction" (Paper delivered at IEEE Int'l Symp. on Robot & Hum. Interactive Comm., 6-8 September 2006) (2006), IEEE Int'l Symp. on Robot & Hum. Interactive Comm. 80 at 83.

Similarly, an AA's perceived autonomy may amplify its responsibility — both with regard to good and bad news.[19] That is, an AA messenger that appears relatively autonomous receives more credit for good outcomes and more blame for poor outcomes than a less autonomous AA[20] — likely because it appears to retain a greater sense of agency in the matter.[21] Such responses may stem from the diffusion of responsibility — individuals have a new person to which to spread credit or blame.[22] Especially if individuals are unfamiliar with technology, they may have higher expectations for an AA's performance compared to reality.[23]

As such, anthropomorphism seems to be the main driver of how humans feel about their interactions with an AA.[24] That said, compared to humans and computer models, robots appear to be viewed differently[25] and more harshly. In a 2 (instructor: human vs. robot) x 3 (news: good, bad or neutral) experimental design, human instructors of a course — and their feedback, regardless of valence — were rated more favorably than agents and their feedback. Moreover, participants who received negative feedback from a robot instructor especially disliked their instructor.[26]

It appears that AA messengers receive the responsibilities — but not the rewards — of agency. At least for now, AAs do not yet have the "higher order mental capacities" of humans.[27] Yet even in exhibiting agency by incorporating

---

[18] See Kerstin Fischer & Anton Batliner, "What Makes Speakers Angry in Human-Computer Conversation" (Paper delivered at Proc. of the Third Workshop on Hum-Computer Conversation, 3-5 July 2000) (2000), Third Workshop on Hum-Computer Conversation 62 at 67. The authors also look at negative responses to misunderstandings and lengthy processing time by the AA. *Ibid* at 62.

[19] See Alexander Serenko, "Are Interface Agents Scapegoats? Attributions of Responsibility in Human-Agent Interaction" (2007) 19:2 Interact. Comput. 293 at 293.

[20] See *ibid*.

[21] See Nathan Stout, Alan R. Dennis & Taylor Michael Wells, "The Buck Stops There: The Impact of Perceived Accountability and Control on the Intention to Delegate to Software Agents" (2014) 6:1 AIS Transactions on Hum-Computer Interaction 1 at 1 (2014). Note that the AA doesn't need to have greater agency, just to *appear* to; recall that participants are more likely to attribute agency to similar agents. See Li et al., *supra* note 15 at 195; Waytz, Cacioppo & Epley, *supra* note 13 at 226–227.

[22] See Kim & Hinds, *supra* note 18 at 80.

[23] See Lee et al., *supra* note 16 at 205.

[24] See Waytz, Cacioppo, & Epley, *supra* note 13 at 226–227.

[25] See Christoph Bartneck, Juliane Reichenbach & Julie Carpenter, "The Carrot and the Stick: The Role of Praise and Punishment in Human-Robot Interaction" (2008) 9:2 Interact. Stud. 179 at 179.

[26] Eunil Park, Ki Joon Kim, & Angel P. del Pobil, "The Effects of a Robot Instructor's Positive Versus Negative Feedbacks on Attraction and Acceptance Towards the Robot in Classroom" in Bilge Mutlu et al., eds., *Social Robotics* (Berlin: Springer, 2011) 135 at 135.

[27] Waytz, Cacioppo, & Epley, *supra* note 13 at 220–221.

its programmers' or customers' preferences, an AA messenger could still be blamed — without any way to "fight back" — for not just execution mistakes but also giving bad news to its clients. In fact, a debate among philosophers of science has emerged on whether AAs should bear responsibility (either for bad news or a mistake), even if they have both agency and accountability.[28] The question of who bears responsibility is one that I will elaborate on in the case studies.

Moreover, it certainly seems possible that an AA's users might confuse it for acting in the "wrong" client's interests. For example, imagine an AA owned by a travel company that is supposed to find the "best deals" for the customers. When the AA is unable to find available fares for a particular trip, the users might think that the travel company has programmed the AA to only show expensive fares and might even sue for fraud. Transparency may help in this regard. In one study, robots that explained their actions as they went — and were therefore transparent — were rated as less blameworthy. However, transparency can compound confusion if the AA does not provide lay explanations.[29] While human messengers deal with issues of transparency (i.e., conflicts of interest) popular intrigue surrounding robots may compound suspicion.

I. IN WHAT SITUATIONS SHOULD AN AA REVEAL BAD NEWS?

While the fact that individuals tend to react more poorly to AAs might initially suggest that only humans should be the messengers of bad news, there are certain cases in which it would be appropriate for AAs to act in such a capacity. First, cost or resource constraints might make it impossible to humans to deliver all bad news. Second, there may be some situations in which it is nearly impossible for a human to share, or to be compelled to share, bad news.

How should we choose which domains are appropriate for AA messengers? Firms and other potential developers of AA messengers might pragmatically choose to use AA messengers in domains in which the damage of AA messengers are minimized. That "damage function" could include the loss in profit from lost customers or the costs of developing a more anthropomorphic AA. On the other hand, there might be some domains in which AA messengers are especially good with bad news — for instance, an individual may want privacy in certain areas.[30] From an ethical standpoint, there are certain areas in which AA messengers appear to be completely inappropriate, such as informing someone of a death in

---

[28] See Judith Simon, "Epistemic Responsibility in Entangled Socio-Technical Systems" (Paper delivered at the AISP/IACAP World Cong, 3 July 2012) (2012) AISP/IACAP World Cong. 49 at 49.

[29] Kim & Hinds, *supra* note 18 at 85.

[30] Consider the United States' Transportation Security Administration's emphasis on the fact that images from early-stage body scanners would be viewed in a separate room. TSA, "Airport Testing of New Advanced Imaging Technology Software Begins Today" (1 February 2011), online: *TSA* (blog) <blog.tsa.gov/2011/02/airport-testing-of-new-advanced-imaging.html>.

his family. Such a message violates the recipient's dignity by saying they are not worth a human messenger's time.[31] Alternatively, perhaps a reliance on automation will relax our collective "bad news muscle," that is, cause individuals to become increasingly uncomfortable and unable to reveal hard truths.

The challenges in choosing appropriate domains are similar to most public policy challenges, although one could argue that the nature of bad news simply necessitates greater sensitivity. Regulator deference is one option, but such an assumption relies on the benevolence of the parties in much the same way that benevolent deception would. Citizen-driven preferences for messengers are another alternative. In one case, surveys could measure citizen preferences and the strength of these preferences and each jurisdiction or firm could adjust its allocation of human messengers accordingly (i.e., a triage process). It could be the case that a minority of citizens has a strong preference for human AAs in one particular domain, in which case the firm or regulator could either choose to keep only human messengers in that domain. Perhaps, the firm or regulator could ask users ahead of time what they would prefer,[32] although the users might have inconsistent preferences.

Certainly, a third way would be for a human to supplement an AA messenger: for example, if a situation escalates (e.g., the receiver becomes extremely emotional) a human might take over the process. But moving from an AA to human messenger could be problematic, particularly if doing so reveals the fact that the AA is disguised.

II. A PROPOSED FRAMEWORK

Before discussing applications, it is important to note that the role of a messenger — especially an AA messenger — may depend on who is responsible for the bad news. There appear to be three classes of bad news which I'll respectively label "class 1," "class 2" and "class 3":

1. bad news for which no one (or a third party) is entirely responsible (e.g., medical results performed correctly by a clinic),
2. bad news for which the messenger is (at least partially) responsible (e.g., medical results performed incorrectly by a clinic), and
3. bad news for which the recipient of the news is (at least partially) responsible (e.g., medical results due to a patient lifestyle).

Depending on the class each situation falls into, however, there will be different effects on liability.

---

[31] Obviously, this depends on how relative respect there is for individuals versus robots, which may change over time.

[32] Such "active choice" systems exist in the retirement savings literature. See generally Gabriel Carroll et al., "Optimal Defaults and Active Decisions" (2009) 124:4 Q J Econ 1639 at 1639—1676.

Generally speaking, if a class 2 situation occurs, the messenger might be liable for the underlying news. Breaking bad news — especially through an AA messenger — might therefore exacerbate the anger of the recipient at the messenger, potentially leading to a greater frequency of claims. Of course, the question is whether a jury, if applicable, might see this through the perspective of messenger or recipient. In contrast, a class 3 situation resembles a whistleblower situation, for which some protective policies exist. In rare situations, messengers of class 1 news could be exposed to litigation risk if their news led to extreme responses.[33]

Incidentally, in all three classes, human messengers are less likely to share news, which might be a good reason for the use of an AA messenger. Yet claims may be more likely with an AA messenger if recipients feel particularly wronged (class 2), defensive (class 3), or more upset (class 1).

III.   Examples

(a)   The role of lawyers

Lawyers are traditionally thought of as counselors or as translators between clients and the workings of the court system.[34] On the other hand, they may also function as "reputational gatekeepers," who might have to reveal evidence of a firm's non-compliance.[35] Because their client may change counsel, lawyers may be hesitant to provide bad news. For example, counsel could blame their poor litigation outcome on their counsel's quality of work even if the facts of the case were not helpful to begin with. Similarly, clients may be dissatisfied by a settlement that their counsel has negotiated for and charge the counsel with being more aggressive than is necessary. Ultimately, in negotiation and trial settings, a client's distrust of her lawyer is likely to be exacerbated in the stressful situation that litigation creates:[36] one scholar points out that "[t]he indemnifying attorney who reluctantly goes to trial pays a high price for following his client's orders. The punitive aspect of such risk allocation is the equivalent of shooting the messenger who bears bad news."[37] That is, breaking bad news not only leads to

---

[33] E.g., an IIED argument. One scholar notes: "The question here really is whether the law should recognize a duty of care to convey bad news tactfully so as to minimize the risk of physical consequences from sudden shock." Calvert MacGruder, "Mental and Emotional Disturbance in the Law of Torts" (1936) 49 Harv L Rev 1033 at 1045-1049.

[34] See Clark D. Cunningham, "A Tale of Two Clients: Thinking About Law as Language" (1989) 87:8 Mich L Rev 2459 at 2465, 2468.

[35] See Andrew F. Tuch, "Multiple Gatekeepers" (2010) 96 Va. L. Rev 1983 at 1595; Pamela S. Karlan, "Discrete and Relational Criminal Representation: The Changing Vision of the Right to Counsel" (1992) 105:3 Harv L Rev 670 at 684.

[36] See generally Sam Margulies, "Representing the Client from Hell: Divorce and the Borderline Client" (1997) 25 J Psych & L 347 at 347-364, for one example.

[37] Bradley L. Smith, "Three Attorney Fee-Shifting Rules and Contingency Fees: Their Impact on Settlement Incentives" (1992) 90:7 Mich L Rev 2154 at 2166.

worse outcomes for the messenger, but also for the receiver, for example, the new counsel may be more unfamiliar with the case overall.

Yet replacing or supplementing lawyers' role with a computer or AA messenger may be problematic. Researchers have found that laboratory participants are more likely to cooperate with humans in a Prisoner's Dilemma scenario and to be fairer in the Ultimatum Game, compared to robots,[38] suggesting that more aggressive negotiating may result if an AA messenger is used to relay negotiations between two opposing parties. Rather, lawyers could more plausibly use AA messengers to show the bad news of low likelihood of trial success. That is, presuming that a machine learning tool exists to determine the probability of cases, an AA messenger could provide a client with the probability surrounding trial success or of expected settlement amounts.[39] Compared to hearing about trial probabilities from an attorney, an AA might appear to be less biased, depending on how transparent (i.e., the AA could show the inputs) or anthropomorphic it appears to be. As such, people's optimism regarding a successful outcome might be better calibrated under an AA. However, an AA messenger might be less persuasive if it is trotted out to the client only when probabilities are low according to the counsel. That is, a client could think that the AA messenger is rigged to tell only bad news and support counsel's name. As a matter of precaution, it might be worth showing the AA messenger more often than not, along with the inputs. Of course, concerns about automated decision support are then more likely to accompany a client's already existing concerns regarding an AA messenger.[40]

(b) Medical diagnoses and genetic disclosure

The existing literature has focused on this category for good reason: breaking bad news in a medical context often involves a single revelation with significant consequences. While all three classes of news might be observed in the medical domain, the legal literature on medicine has focused on class 2 cases, in which the messenger has erred. As expected, there is a gap between how much people claim they would report a medical error and actual levels of reporting.[41] Moreover, reporting levels appeared to decrease after hospitals implemented increased liability for disclosure,[42] suggesting that medical professionals preferred to take the chance of heavier penalties and lower probabilities of penalty, with broader

---

[38] Eduardo Benitez Sandoval et al., "Reciprocity in Human-Robot Interaction: A Quantitative Approach Through the Prisoner's Dilemma and the Ultimatum Game" (2015) 8:2 Intl J Soc Robotics 1 at 1.

[39] See Harry Surden, Note, "Machine Learning and Law" (2014) 89:1 Wash L Rev 87 at 104.

[40] Consider the resistance of some judges in implementing sentencing guidelines. See e.g. Alison Siegler, "Rebellion: The Courts of Appeals' Latest Anti-Booker Backlash" (2015) 82:1 U Chi L Rev 201.

[41] Lauris C. Kaldjian et al., "Reporting Medical Errors to Improve Patient Safety" (2008) 168:1 Archives of Internal Med. 40 at 40.

implications for medical care overall.[43] Here, an AA messenger could "force" increased disclosure if there were an automated system that could flag discrepancies in data; the AA could automatically threaten to break the news in cases where an error likely occurred. That could even trigger a human messenger to speak up and break the news, particularly if the human messenger was afraid that the AA messenger would be perceived poorly (for breaking the news) and that they themselves would be perceived especially poorly (for letting the AA do so or for seemingly covering up the error). However, the consequences of a false positive error would be severe for medical professionals and they would likely resist this development.[44]

Understanding how bad news breaks — and the role of AAs in this matter — will be especially important as patients increasingly seek and rely on genetic results. The idea of returning genetic results, either from individual testing or large-scale research studies is daunting for researchers, due to potential class 2 issues of faulty disclosure (e.g., premature disclosure or disclosure of inaccurate results), but also consequences from breaking the bad news itself. For example, a patient might sue based on negative consequences resulting from their actions in reliance on otherwise accurate genetic results (e.g., pursuing a particular speculative therapy).[45] Such a result might occur regardless of an AA or human messenger, but AA messengers could be shown to increase the propensity of these actions.[46]

One could also imagine a situation in which the effects of an AA messenger might be alleviated: when patients go to the clinic for serious medical issues, they may be accompanied by a friend or family member. In those cases, an AA messenger might be a good intermediary, one that delivers the news to the friend or family member, who in turn delivers it to the patient. Such an approach considers the dignity of the patient, and minimizes the pain of the medical professional. While the friend or family member might face the pain of breaking the news,[47] they could consent to taking that responsibility ahead of time. The

[42] Allen Kachalia et al., "Liability Claims and Costs Before and After Implementation of a Medical Error Disclosure Program" (2010) 153:4 Annals of Internal Med 213 at 213.

[43] See William M. Sage, "Regulating Through Information: Disclosure Laws and American Health Care" (1999) 99:7 Colum L Rev 1701 at 1794.

[44] In South Africa, the rule of "therapeutic privilege" allows doctors to withhold information from patients if the doctor believes revelation of this medical information would not be in the patient's best interest. L.C. Coetzee, "Medical Therapeutic Privilege, a Separate and Independent Defence Eo Nomine?" (2004) J S Afr L 464 at 464.

[45] See generally Heather Skirton, "Direct to Consumer Testing in Reproductive Contexts — Should Health Professionals Be Concerned?" (2015) 11:4 Life Sci, Soc & Pol'y 1 at 1-9.

[46] For example, patients may be more likely to engage in risky decisions, given a higher increase in negative emotions caused by an AA messenger, or a disregard for the AA's advice, as we have discussed. To be sure, this is an empirical question.

[47] Friends and family members may be less sensitive to the bad news than most people believe. See Kate Sweeny, James A. Shepperd, & Patrick J. Carroll, "Expectations for Others' Outcomes: Do People Display Compassionate Bracing?" (2009) 35:2 Pers & Soc

only legal questions that might remain are whether the friend or family member is able to accurately convey the AA messenger's results and who would be liable for any errors as a result of this handoff problem.

IV. WHAT IS THE FUTURE OF BREAKING BAD NEWS?

I have tried to highlight some of the legal issues surrounding the use of AA messengers in breaking bad news. In doing so, I have also tried to provide background on existing research surrounding the breaking of bad news with regard to human and AA messengers — both understudied phenomena. The examples in this paper are by no means exhaustive. Consider a few more: could an AA messenger increase communication between police and victims in a helpful way? Could an AA messenger help to decrease groupthink by providing a perspective whose ex ante principles everyone agrees on? Context appears to be key and further experimentation or case studies will point to some generalizing principles.

Specifically, future research might focus on how people respond to AA messengers in each of the three classes of my framework. What we know now about messengers is often entwined with and dependent on the party that made the error in the first place. There is a great deal of legal research involving the role of intermediaries.[48] Here, intermediaries face the question of whether they have a duty to disclose or remove information. In contrast, an AA messenger's actions per se are not alleged to cause a wrong; instead its automated nature is the offense. Considering the implications from this new kind of intermediary will provide a unique extension to this literature. Moreover, the literature on intermediaries focuses on the liabilities of a firm due to the acts of a third party. In this paper, I have assumed for now that the firm would be one that enlists an AA (i.e., the third party) for its customers, but if a firm contracts out its messaging to AAs owned or designed by another company, there will certainly be disputes over who is liable for the AA.

More broadly, what might the automation of bad news reveal for AI-human communication? At the very least, the breaking of bad news is an edge case: because bad news often encroaches on sensitive issues, the worst case scenarios (i.e., responses to bad news) are likely to be among the worst of all AI-human communication scenarios. But even further, current concerns regarding AI-human interaction concern issues of dignity and autonomy — things we consider to be fundamental to a well-lived life. What does dignity — or consideration for

---

Psych Bull 160 at 160 (showing that "participants braced [their emotions in response to bad news] for friends only when the friend's performance had self-relevant implications").

[48] See e.g. Mark Bartholomew "Cops, Robbers, and Search Engines: The Questionable Role of Criminal Law in Contributory Infringement Doctrine" (2009) BYU L Rev 783 at 822; Daryl J. Levinson, "Aimster and Optimal Targeting" (2007) 120:5 Harv L Rev 1148 at 1153.

another person's well being look like? Should we fear the increasing power of technology — or worry about the people who might be able to control it? How is the role of an AA designer different from a human whose interactions are based on simply wanting to be liked (or a manipulative person, in the worst case)? The study of breaking bad news not only foregrounds these issues in the spotlight, but also suggests policy solutions that may be applicable to other aspects of the AI-human relationship.