



Outsourcing Emotions to GenAI: The Risk of Bypassing Identity Development

Ten-Herng Lai¹ · Alfred Archer²

Received: 5 June 2026 / Accepted: 16 September 2026
© The Author(s) 2026

Abstract

There is growing research on how and when the use of Generative AI (henceforth GenAI) in cognitive tasks may lead to a decline in the users' cognitive capacities. There is emerging evidence that Generative AI also presents us with the opportunity to outsource some emotional tasks, such as flirting on dating apps or apologising. We believe that this may enable unsuspecting users to outsource too much of their affective capacities. Here, we wish to focus on a particular problem: insofar as articulating our own emotions is essential to discovering what we value and thereby shaping what sorts of persons we are, users who are heavily dependent on GenAI for emotional tasks risk forgoing crucial opportunities to develop a proper sense of their own identity.

Keywords AI · Deskilling · Affective injustice · Emotions · Outsourcing · Capacity-hostile environment · Affective resources and opportunities

1 Introduction

In “Deep Learning,” an episode of the American television comedy show *South Park* (Parker, 2023), a group of schoolboys start using Generative Artificial Intelligence (henceforth GenAI) to message their girlfriends. The girls are quickly impressed by how emotionally mature and sensitive their boyfriends have become, as the messages they are receiving change from a simple thumbs-up to elaborate emotionally validat-

✉ Ten-Herng Lai
laitenherng@gmail.com

Alfred Archer
a.t.m.archer@tilburguniversity.edu

¹ University of Stirling, Stirling, UK

² Tilburg University, Tilburg, The Netherlands

ing responses. As this fictional depiction illustrates, we can outsource many emotionally difficult or loaded tasks to GenAI. But should we?

Consider apologising. Apologising isn't easy. It involves working through our problematic past, seeing from the perspective of the wronged party how we have wronged them, feeling guilt and remorse, and putting all of it into words, not to mention making a promise that we will be better (Tobi, 2025). But that's a lot of information and unpleasant feelings to go through. So we might be tempted to act like pseudonym Avery's boyfriend, who used to have "a tendency to overlook things written in texts" and keeps ignoring her feelings (Tait, 2024). When Avery sent him another long text explaining her feelings and asking her boyfriend to carefully read through it before responding, 15 min later, she received a long text validating all her feelings. It felt great, until it didn't. Her boyfriend didn't write any of it. ChatGPT did. The boyfriend didn't work through any of the information, but treated the apology like a box that could easily be ticked by GenAI. Not *working through* anything at crucial moments like this is what we are primarily concerned with in this paper.

There are many emotionally complex and difficult tasks, and there is much temptation to outsource. GenAI not only has a rich vocabulary and avoids grammatical errors, but can also spare us from the emotionally taxing tasks of confronting difficult memories, making sense of all the conflicting and unorganised raw feelings, facing our own misdeeds, or any of the unpleasantness. But outsourcing the emotional labour to GenAI has significant drawbacks: deception, the lack of sincerity, and, of course, the worry that we won't remember any of it and expose ourselves when we eventually interact face to face with the recipient. Moreover, what GenAI writes draws from training data rather than the specific persons our emotions are about, and thus lacks the correct relation with the persons we interact with (Weirich & Holdier, 2026). All these issues and the wrong done to those on the receiving end are well known. Here, we propose that there's a further important problem: outsourcing difficult emotional tasks to GenAI risks impoverishing *the users'* own affective capacities. We believe that being emotionally impoverished can be bad for us for various reasons. Here, we wish to focus on a particular setback: Insofar as working through our emotions reveals to oneself what one values, bypassing such moments amounts to foregoing opportunities to develop or discover one's identity.

Here's the plan. Section 2 introduces the existing literature on the well-known *cognitive* impact of outsourcing cognitive tasks to GenAI, and elaborates on when outsourcing risks problematic deskilling. In Sect. 3, we outline the idea that affective resources and opportunities contribute to the development of our affective capacities, and that engaging in the task of interpreting and expressing our emotions is key to understanding who we are and what we value. Section 4 argues that improper uses of GenAI may deprive users of the opportunity to work through their emotions, particularly when users completely outsource the articulation of emotions to GenAI or when they use GenAI to compose initial responses. Section 5 introduces the concept of *capacity-hostile environments* – settings where we are severely tempted to forgo activities that maintain certain important skills, and argues that the ready availability of GenAI creates pervasive capacity-hostile environments in our everyday lives in relation to our affective capacities. Section 6 distinguishes problematic GenAI emotional deskilling from curating other affective artefacts (such as get-well cards) to

express our emotions. The conclusion briefly spells out a structural issue that makes the issue at hand difficult to resolve.

2 GenAI and Cognitive Outsourcing

There is a growing literature on how the use of GenAI in writing, especially when it comes to assignments, can create problems for learners. Outsourcing the reading and writing to GenAI can leave users with impoverished critical thinking skills (Gerlich, 2025), poor argumentation skills (Stadler et al., 2024), lower levels of creativity and independent thinking (Zhang & Xu, 2025), and a decline in memory retention (Bai et al., 2023), just to name a few. The basic idea is simple. When people outsource cognitive tasks to GenAI, brain areas typically associated with the relevant tasks are no longer active (Kosmyna et al., 2025). Outsourcing, therefore, risks forgoing opportunities to properly use and develop the relevant cognitive capacities.

We are well aware that not all cases of cognitive outsourcing are problematic. We have largely outsourced, for example, much of our basic mathematical tasks to calculators, and such outsourcing appears to have reduced our mental load so that our time and attention can be put to higher-order activities. Such outsourcing appears to be largely innocuous, as long as it is not done in phases where we acquire important basic skills that are crucial to acquiring more advanced skills. Yet, there are important differences between using calculators to calculate and using GenAI to write. Relevant to the purpose of this paper, we will follow Buijsman et al. (2025) and introduce two important problems associated with the (over)reliance on GenAI.

First, outsourcing becomes problematic when GenAI is largely reliable but can nevertheless make mistakes, and yet the mistakes are often difficult to detect. When mistakes are made, the output no longer advances the users' goal(s), but the users are often unaware of such problems. Compare this to calculators. Calculators are highly reliable, and in the rare cases they malfunction, the problem is easily detectable, e.g. not displaying an answer at all or displaying an error message. Users can thereby easily avoid accepting erroneous answers provided by a malfunctioning calculator. In contrast, when GenAI produces hallucinations or introduces biases, it largely still presents itself as confident and authoritative. And here, it is up to the users to detect those mistakes.

Yet, the ability to detect those mistakes can be eroded by the continued reliance on GenAI. When we completely outsource tasks to GenAI, we simply lack the developed skills to work out what correct answers look like. Even in the somewhat less egregious cases where the human user plays a "supervisory" role, namely, allowing GenAI to produce initial drafts and then have the user check and endorse the output, things are hardly better. As most of the "hard work" is offloaded to AI, the human user becomes reliant and tends to uncritically accept AI's output. For example, one study (Candelon et al., 2023) found that for tasks beyond the strength of GenAI, collaborating with AI leads to inferior outputs. In their study, when engaging in business-solving tasks, consultants who use GPT-4 to write initial drafts produce inferior reports compared to those who write independently. Similarly, when human users make decisions in collaboration with AI, their ability to make independent decisions

in the long run is also compromised, as demonstrated in domains such as financial services (Wessel, 2023) and medical diagnosis (Goddard et al., 2011, 2014). In such cases, human users tend to become further dependent on AI.

Second, GenAI is known to embed certain biases (Lin & Chen, 2022; Lin & Huang, 2026). When we are inclined to outsource important tasks to AI, we risk being affected by these biases. Even if we work to critically engage with the output of AI, by collaborating with the machine, we risk not only being exposed to such biases, but adopting them (Buijsman et al., 2025). For example, in an experiment (Jakesch et al., 2023) that asks subjects to write opinion pieces on whether social media is overall beneficial to the society, subjects “assisted” by GenAI were more likely to produce pieces that reflect the biases of the AI model (which were determined by the experimenters). Most concerningly, subjects later held onto the biases of GenAI, and were predominantly unaware that the AI in question was biased, even if the biases are in conflict with the subjects’ own prior inclinations (see also Kidd and Birhane (2023), Vicente & Matute, 2023). A further study indicates that people can be affected even if warned against GenAI biases (Williams-Ceci et al., 2026).

Having explored these two problems, let’s now turn to the motivation for resorting to AI. Drawing from academic cases, students resort to GenAI cheating for various known reasons. Here are a few. First, students feel the pressure to acquire decent or at least passing grades (Cong-Lem et al., 2024). Second, students may find approaching deadlines daunting and resort to quicker ways to complete the assignment (Kapoor et al., 2025). Third, while the majority of students understand that GenAI risks compromising their skill development (Chan, 2024; Pitts et al., 2025), a minority of students perceive no risk whatsoever in GenAI cheating (almost 5% in Pitts et al., 2025: 5), and *ipso facto* have not considered the negative impact on their learning. The risks and motivations of cognitive outsourcing will become relevant in later discussions, as we believe that there are many commonalities with outsourcing emotional tasks.

To briefly recap, here’s what we know. Certain ways of using AI are problematic. These cases involve outsourcing the tasks completely to GenAI, as well as uses that involve letting GenAI do most of the heavy lifting, especially in crafting the first draft. In such cases, the users risk deskilling on one hand, and are unable to detect whether outputs are problematic due to the deskilling on the other. Collaborating with GenAI also risks adopting the biases of the AI model. Yet, there is still much temptation to outsource: when one is working under the pressure of deadlines, where stakes are high, sometimes compounded by not knowing the consequences of deskilling. These considerations will become relevant when we move on to the problems with outsourcing emotional tasks.

3 Affective Resources and Opportunities

We are affective beings, and one of the ways in which we may be wronged by other people is in relation to our affective responses (Srinivasan, 2018). For instance, Francisco Gallegos (2022) proposed that we can be wronged as affective beings by having affective resources and opportunities withheld from us. For example, being excluded from educational resources and opportunities that hone one’s affective capacity can

render one affectively deprived (195). By lacking certain concepts, we may be left in a position where we are unable to make sense of our emotional experiences (196). We may be left confused by our own emotional responses, unable to work through trauma, unaware that we need to seek help, or become emotionally desensitised or hypersensitive to unfamiliar situations and thereby interact with others in disrespectful and harmful manners. For example, lacking the concept of post-natal depression may render some unable to properly understand their own experience (see Fricker, 2007: 149 for how this constitutes a hermeneutical gap) and risk being dismissed as an unloving and incompetent mother when not displaying sufficient affection to one's offspring. Having the concept, on the other hand, helps people in need seek help, and receive the much-needed emotional and physical support (which enables most to recover).

Our proposal is that alongside such resources and opportunities, we also have affective abilities or skills that can be developed and improved or that may be left undeveloped, and the development has much to do with the people we interact with. An important skill is the ability to interpret and make sense of our own affective experiences. This is an important skill, as emotional interpretation and expression play a crucial role in developing a sense of who we are and what we value. Talbot Brewer (2011, 286) argues that emotions play an important role in revealing to us what our pre-reflective evaluations of the world are. Suppose Eileen thinks of herself as someone who is not very ambitious and for whom their work plays little role in their identity. Now imagine that Eileen is promoted and feels proud as a result. These feelings of pride may indicate to Eileen that this form of recognition is more important to her than she had previously realised. In this way, our emotions can reveal how we value the world in ways we may be unaware of.

Moreover, through interpreting and expressing our emotions and trying to make sense of them to ourselves and others, we come to a more refined articulation of who we are and what we care about (Brewer, 2011, p. 287). For instance, as Eileen tries to make sense of her feelings of pride and explain them to her close friend, she is likely to revise her sense of what matters to her and come to articulate in a new way what it is that is important to her. This process of interpreting our own emotions and articulating them to others plays a crucial role in shaping our identities and our values. By making sense of our initial affective responses, thinking carefully about what they reveal about who we are and articulating this to others, we come to a more refined understanding of who we are and what we value. The process of emotional articulation is of crucial importance here. What Brewer calls the process of bringing "one's half-formed evaluative sensibility into words" (Brewer, 2011, p. 289), plays a crucial role in pushing us to find a way of bringing together our conscious sense of who we are with our less conscious, pre-reflective evaluations of the world.

When we do not engage in such a process of self-articulation, there is a risk that we will become alienated from our emotions. Brewer (2011, 275) identifies two kinds of risk here. First, we may find ourselves expressing an emotional outlook that is not our own. For instance, if Eileen expresses anxiety about her promotion because she thinks that this is what her partner would expect and not because she actually feels anxious, then she is alienated from her emotions in this first sense. Second, we may be alienated from our emotions when we refuse to accept the evaluative

outlook that the emotions contain as being our own. For instance, if Eileen is so intent on expressing this anxiety that she refuses to interrogate her half-formed feelings of pride about her promotion, then she is alienated from her emotions in this second sense. As Brewer (2011) puts it in relation to the example he focuses on, this runs the risk that her “inchoate feelings are left in crude form, deprived of that extra degree of intelligibility and determinacy that might have been conferred upon them by a sustained effort to articulate the potential goods they intimated” (295). In other words, Eileen will be closed off from a proper process of determining how she really feels about the promotion and what these initial feelings may reveal about what she truly values. While we cannot, of course, pursue all possible opportunities for self-knowledge and articulating our feelings to ourselves and others, we take it that to habitually refrain from doing so throughout one’s life would lead to an impoverished understanding of the self.

While Brewer’s focus is on the ways in which we make sense of our own emotions, it is important to note, as we have already hinted at, that this is often a very social process. As Sue Campbell has argued, the process of communicating one’s emotions to others and receiving their feedback plays an important role in shaping our feelings (1997, 76). When others respond dismissively to our emotions, this can make it more difficult to articulate to ourselves what it is that we are feeling. For example, if Eileen were to express her pride to her friend and her friend were to respond by telling her “Don’t be silly, you don’t care about work,” then this may lead her to be less sure about what it was she was feeling (Campbell, 1997, p. 110). On the other hand, if Eileen’s friend were to respond in a supportive way that encouraged Eileen to articulate more precisely what it was she was feeling, then this may help Eileen to come to a more precise understanding of how to feel about her promotion and what role she wants success in the workplace to play in her identity.

The ability to act as what Sue Campbell calls a “sympathetic interpreter” (1997, 165) – someone who is able to support others in coming to a more articulate understanding and interpretation of their feelings – is something we may be more or less skilled in. Effective interpreters will not be those who will always simply affirm our first attempt to make sense of our feelings. Rather, they will be people who help us get to a more refined understanding of what it is that we are feeling. While some people are incredibly skilled at helping others determine what it is they are feeling, other people lack these emotional skills. Many of us will know the friends or family members we would go to when we want help understanding how exactly we feel about something, as well as those we would never think of discussing these issues with.¹ Having friends who can fulfil the role of a sympathetic interpreter is highly beneficial for us to properly understand what we genuinely care about. Those who do not know anyone who can help them interpret and articulate their emotions in a helpful way will find it more difficult to come to a clear understanding of how they feel and what they value.

¹ Campbell (1997, Ch.5) also emphasizes that this is not only an individual skill but one that can be developed in interpretative communities, particularly in subcultures that reject dominant norms for emotional expression and interpretation.

So far, there are a few things worth highlighting. First, in certain crucial moments, our emotions, along with our effort to work through them, can reveal how we value the world in ways we may be unaware of. If, somehow, when these crucial moments occur, we do not take them in, our emotions would not be able to reveal anything to us. Second, even if we do feel something during such crucial moments, we may not be able to immediately fully process our emotions. In such times, we can work to articulate our feelings. Sometimes, we also consult others, and depending on whether those around us can act as sympathetic interpreters, our understanding of our own values can be bolstered or hindered. Both of these will become relevant when we later discuss the impact of GenAI on our emotions and, in turn, our self-understanding.

Now, one's opportunities to make sense of their own emotions aren't just shaped by the people around them. Recent scholarship has documented how the *artefacts* in the built environment can affect one's emotions (e.g. Krueger, 2023). These artefacts may be designed to promote the interests of those affected. For instance, a beautiful work of art in a public place may promote positive feelings in those who encounter it and help to improve their mood. However, affective artefacts can also constitute forms of *hostile scaffolding* (Timms & Spurrett, 2023), that is, features of the environment that scaffold certain affective responses that undermine the interests of those who are scaffolded whilst at the same time promoting the interests of those doing the scaffolding. Playing classical music, for example, can make a physical location less welcoming to teenagers (Hirsch, 2007; Osler et al., 2025). Commemorations such as statues and monuments can also demand that visitors be "respectfully silent" towards symbols of oppression (Lai, 2025) or make the city space more homely to colonisers at the expense of the colonised (Archer, 2024).

We contend that we should also understand GenAI as a form of affective artefact, that is, things created or modified to scaffold affective experiences (Piredda, 2020). GenAI can directly affect our emotions (as in cases of AI companions) or, as per the focus of our paper, create an environment hostile to the *development* of affective capacities. We believe that GenAI services run the risk of prompting us to habitually forgo opportunities to develop these skills. We think this is highly related to how the inappropriate use of GenAI leads to cognitive outsourcing, and risks leaving users who are highly reliant on such technologies cognitively deprived, accumulating a cognitive debt, so to speak.

4 Affective Artifacts and Emotional Articulation

In this section, we shall examine how GenAI can be used to help facilitate (or be brought into) the process of articulating one's own emotions. We will explore how regular and habitual uses of GenAI for the purposes of emotional articulation can be deeply problematic. This is particularly so when this involves the complete bypassing of our emotions: through not confronting our emotions, we may constantly forgo opportunities to discover or come to work out what we genuinely value. We will also show that even in cases where there is some human oversight – where the user reads, edits, and approves the content – we may end up bypassing the difficult work of emotional articulation that we, following Brewer, have claimed is valuable. We will then

contrast the most problematic uses with cases where GenAI plays a more supportive role. While we think in such cases, GenAI can often help us better understand our emotions, there is nevertheless a slight risk of allowing GenAI undue influence on our self-understanding. (We will leave the worry that GenAI can lead to *widespread* problematic uses to Sect. 5).

Here's what we already know. There are websites purporting to provide GenAI to complete specific emotional tasks: interpreting one's own emotions, writing a breakup letter, a eulogy, a love letter, an apology, and so on. The AI companion Pallie (n.d), for example, offers a Breakup Text Generator that promises to provide a "kind, clear and respectful" breakup message that people can send to those they wish to break up with. Of course, one does not need to make use of any specialised service here, one can simply ask a popular LLM like ChatGPT or Gemini to write a breakup text or other text designed to express a particularly emotionally laden message.

For the sake of argumentative simplicity, we will distinguish three types of users of GenAI for emotional tasks.

Complete Outsourcers These users outsource the emotionally expressive task entirely to GenAI and use the text it has created without critically reflecting on whether it provides a reasonable reflection of how they feel.

First Drafters These users start with a rough idea of what they feel and ask GenAI to create a draft text that expresses their feelings. They then check and, if need be, edit the text to ensure it is (or appears to be) a reasonable reflection of how they feel.

Reflective Engagers These users will use GenAI as a tool to help them reflect critically on what it is that they are feeling and as an aid in coming to a more refined and articulate account of what they feel that they can then use to inform the emotionally expressive task they are engaging in. These users may specifically ask GenAI to help them establish how exactly they feel about something, for instance through generating questions aimed at facilitating better self-understanding. These users may also work with GenAI by describing how they actually feel, and ask GenAI for concepts and vocabularies to understand and express their feelings.

We recognise, of course, that these are simplified categories and that the boundaries between these different kinds of GenAI users are not going to be rigid. Moreover, someone may fit into one category for some emotional expressive tasks and another category for different emotionally expressive tasks. Nevertheless, we think these categories of users provide a useful starting point for assessing the different ways in which someone might use GenAI for the task of emotional articulation.

Let's start by considering *Complete Outsourcers*. These users effectively use the GenAI tool as a substitute for agency. Rather than performing the task themselves, these users get GenAI to do it for them. In doing so, they completely avoid engaging in the process of trying to articulate to themselves and others what it is that they are feeling. As we explained in Sect. 3, articulating what it is that we are feeling is a central way in which we make sense of who we are and what it is we value. By outsourc-

ing this task, one is effectively giving up on this sense-making process of shaping our identities and our values through the process of interpreting our own emotions.

In doing so, one also runs the risk that one will alienate oneself from one's emotions in either of the ways we considered in the previous section. First, a complete outsourcer may well find themselves expressing an emotional outlook that is not their own. For example, a complete outsourcer who uses GenAI to write a breakup message may send a message expressing gratitude for the time they have spent together as a couple, when they in fact feel only regret for the time they have wasted in the relationship and for not deciding to end things earlier. Of course, there may be cases where a complete outsourcer by chance does happen to express emotions that line up roughly with how they actually feel, but this will be purely down to a happy coincidence.

Second, complete outsourcers close themselves off from their own evaluative outlook that can be found in their underdeveloped and unarticulated feelings. The complete outsourcer who gets GenAI to write their breakup message passes up an opportunity to explore why they really feel that they can no longer continue in the relationship. In doing so, they miss out on a chance to articulate for themselves what does not feel right about this relationship. This could in turn be a useful source of insight into what they value in a relationship and so help guide them in what they should look for in future relationships.

First Drafters are in a better position with regard to emotional alienation than the complete outsourcers. Unlike the complete outsourcers, the first drafters do not completely abandon their agency over the emotional task. Rather, they check whether the message provides a roughly accurate account of how they feel and, ideally, will edit the message when needed to bring it more in line with their feelings if need be. If the first text suggested by GenAI expresses regret when they feel none, then they will be sure to remove this from the message. When first drafters check the message carefully in this way, they are less likely to experience the first form of emotional alienation: that of expressing emotions that are not their own.

However, by outsourcing the first attempt to properly articulate their raw, unrefined feelings, first drafters run a significant risk of closing themselves off from engaging in the difficult task of properly articulating exactly how they feel. By doing so, they close themselves off from the potential insights into their own values that might be gained from this process. Here, they close themselves off from a deeper understanding of their own evaluative outlook that gave rise to these feelings by outsourcing the process of giving more substantial shape and determinacy to them. Consider someone who feels genuine remorse and uses GenAI to write an apology. In this case, the expression may roughly match the person's feelings of remorse. However, the specific way in which the apology expresses remorse will articulate a form of remorse with a particular tone, colour and feel. Influenced by this, the first drafter may indeed come to feel a remorse that matches quite closely with this particular expression of it. But they too run the risk of missing out on important forms of self-knowledge that might be derived from trying to articulate for themselves exactly what they feel. In doing so, they may also outsource the task of providing a more determinate picture of who they are and what they value, both to themselves and to others.

Moreover, first drafters may risk uncritically accepting what GenAI produces, and come to believe that what GenAI produces is indeed exactly what they feel. This worry stems from the observation that in cognitive tasks, users readily accept the first drafts produced by GenAI as correct and reliable. As we discussed in Sect. 2, in different domains (such as performing business-solving tasks (Candelon et al., 2023), using GenAI to produce the first draft forgoes the skill-developing moments, such that users no longer possess the skills to distinguish whether the output of GenAI indeed advances their own goals or is of decent quality. In cases where first drafters overrely on GenAI to produce the first draft, they may be led astray by GenAI in believing that they experience emotions they don't in fact experience, and further foreclose the opportunity to have an accurate understanding of themselves.

In contrast, *reflective engagers* seem largely well-placed to avoid both forms of emotional alienation. Like the first drafters, they would not pass on a message unless they felt that it provided a reasonably accurate articulation of how they feel. They have no special reason to worry, then, about expressing emotions that are not their own. Unlike both the complete outsourcers and the first drafters, though, reflective engagers have no special reason to worry about giving up on possibilities to engage in the process of emotional articulation for themselves. For instance, suppose someone asks GenAI to help them articulate for themselves how they feel about their relationship. The GenAI tool may provide them with a list of questions to answer to try and provide a clearer understanding of their own feelings. When describing how one feels, it is also possible that GenAI brings important affective resources to the user's attention, which may include important concepts that lift the user from certain hermeneutical gaps and even helplines when the users are in desperate need. This could be part of a useful process of giving greater specificity and substance to what it is that they feel that allows them to come to a more refined understanding of their own values. Here GenAI may be used as an aid to sensemaking, rather than a replacement for it. Whether it plays a useful role here or not will depend upon whether GenAI tools are able to function effectively as sympathetic interpreters in Campbell's sense or not. There seems no reason at this stage, though, to rule out the possibility that GenAI tools could function as sympathetic interpreters that can help people using them in the right kind of ways to arrive at a more refined articulation and understanding of how they feel.

There is one minor issue the reflective engager may encounter. We do not wish to overplay this worry, as the reflective engager most likely has significantly better developed affective skills. Yet, we know that sometimes GenAI tools are biased, and when interacting with GenAI, for instance, when seeking advice and feedback, it is possible that the user adopts the biases of GenAI. When this occurs, the user undergoes a shift in fundamental values due to adopting those biases, and such changes are largely undetectable to the user, largely because the embedded biases are hard to detect in the first place (Buijsman et al., 2025). Here, the worry is that whether the reflective engager can advance their own goals through interacting with GenAI partly depends on whether the biases built into the training data of GenAI hinder the process of self-articulation.

Before we move on, it is worth mentioning one additional complication here. We have distinguished between different kinds of users in terms of how they use a GenAI

tool for emotionally expressive tasks. As well as how the tool is used, it is also worth distinguishing between different extents to which a user engages with a particular tool. Some may be habitual users, who use GenAI whenever they have to engage in an emotionally expressive task. Others will be infrequent users, who use it very occasionally, perhaps for tasks they are finding particularly challenging. Both of these different extents to which the tool is used could co-exist with any of the three kinds of user we have mentioned so far. A habitual user could be a complete outsourcer, or a reflective engager. The extent of the use is likely, though, to impact on the extent to which emotional alienation is a problem. The more habitual the user, the bigger any problem of emotional alienation is likely to be.

5 AI and Emotional-Capacity-Hostile Environment

We have, so far, introduced different ways of interacting with GenAI, and how certain habitual uses of GenAI can hinder the realisation and development of our sense of what we deeply value. But how serious is the issue at hand? If only a minority of human beings will opt for the more problematic interactions with GenAI, despite the severe adverse consequences the few would endure, emotional outsourcing would be hardly worth our attention. Yet, we believe that the current state of the art of GenAI and its likely trajectory will entrap more to become habitual outsourcers or first drafters in many different domains, including for emotional tasks. To see this, we will first briefly revisit the motivation for using GenAI to outsource academic tasks, and show that for many, similar motivations hold for emotional tasks. We will then discuss the structural aspect of the problem. We will introduce the concept of *capacity-hostile environments* proposed by Avigail Ferdman (2026), where GenAI can create an environment that effectively invites us to forgo opportunities to engage in emotional tasks core to emotional maturity.

To reiterate, as discussed in Sect. 2, students resort to GenAI because they want to perform decently when the stakes are high, work under the pressure of impending deadlines, and are sometimes unaware of the adverse consequences of cognitive outsourcing. Now consider one use of GenAI: online dating. How is online dating similar to AI cheating? Online dating is hard and frustrating for many. Sustaining a good online conversation takes time, skill, and effort. The more desperate one is to find a proper date (or even *the* proper date as one may see it), the more one may worry about one's own online conversational skills. And where there is a need, there is a market. Academics have recently called for guardrails and regulatory clarity regarding the use of GenAI (Boyle, 2025). The call rightfully highlights concerns about authenticity, privacy, the risk of deception (that may compromise the validity of consent), creating unreasonable body standards, and environmental problems associated with energy consumption. Here, core to our paper, we wish to again highlight the dangers of using GenAI to come up with attractive and interesting things to say during online conversations.

Successful dating, we contend, involves paying attention to the emotional cues of one's counterpart. It also involves paying attention to what they say, and providing appropriate emotional responses to further sustain the interaction. Attention to such

details and being able to provide timely feedback, however, happens to be the forte of GenAI. Such cognitive and emotional tasks can thus be easily outsourced to GenAI. While the output appears to work to a certain extent, the cognitive and affective capacities are bypassed. This outsourcing leaves the user, for the very least, deprived of face-to-face affective skills.² What's more, however, is that we come to know ourselves better when we enter into meaningful relations with others. The emotions we feel when our counterpart(s) respond to us reveal what we value, including but not limited to what we care about in our relationships, what sorts of things we are willing to put effort into, and our bottom lines and dealbreakers, if any. Online dating, thus, presents itself as a scenario where stakes are high (especially for those desperate to connect to others), deadlines are looming (as the need to provide an instant response is there), and where people know little about the consequences of deskilling.

We contend that online dating is but one situation where GenAI appears to be the obvious solution to unsuspecting users. But generally, we believe that GenAI as an affective artefact can create what Ferdman (2026) calls a *capacity-hostile environment*. The basic idea of a capacity-hostile environment is that environments differ in “how favourable they are to capacity cultivation” (3005), and hostile environments, as the name suggests, “restrict, limit or create a narrow field of affordances for capacity development and exercise” (3005). *Affordances*, in turn, are not mere opportunities, but compelling invitations to certain options. For example, a library is a setting where there are not only opportunities to read books, but, given its setting, we perceive the library to be an inviting environment where the setting invites us to read books (Ferdman, 2026: 3004). In this sense, the library provides the affordances of reading, and the library is thereby a *capacity-conducive environment* when it comes to the capacity for reading.

In contrast, having GenAI readily available can transform a setting into a *capacity-hostile environment*. The call for guardrails when it comes to online dating mentioned above (Boyle, 2025) is a response to Tinder (and other dating platforms) making GenAI a readily available feature to assist “struggling users.” When the motivation to take the easier way already exists, making GenAI just a few clicks away isn't merely providing an additional option to users, but can constitute a compelling invitation to resort to GenAI. When the temptation is made salient, users are more likely to accept the offer to outsource, and thereby forgo the opportunity to work out one's own emotions.

The problem we wish to highlight is that the ready availability of GenAI isn't limited to online dating platforms. GenAI doesn't exist in some obscure corner of the internet. The very opposite: the invitation to use GenAI is pervasive. The “AI solutions” to many different tasks are placed at difficult-to-ignore locations on the user interface of the different applications we use, e.g. “Ask Gemini” and “Ask Meta AI or search” appear readily to us when we interact with the services provided by Google and Meta. Google's Gemini, at the time of writing, provides a 12-month free trial for

² Unless all parties put on smart glasses and read the real-time scripts when interacting. But imagine two people focusing on reciting AI-generated text from their smart glasses to each other without really paying attention to the person they interact with. Maybe we can call this *Chinese Room Dating*. We believe that even if this possible is extreme, considering what's morally problematic about it may be instructive.

university students, and some universities actively encourage students to use Microsoft Copilot to take notes and to process reading material. It would seem that such commercial strategies are working. People are increasingly using GenAI as artificial personal assistants to plan their everyday life (Ferdman, 2026: 3007). This leads to our core worry. The capacity-hostile environment created by the ready availability of GenAI is becoming pervasive. Whenever we interact with others online, there will be an invitation to use GenAI. And indeed, not all will eventually come to outsource the majority of emotional tasks, but when the social environment turns hostile, we cannot reasonably expect the majority of users to continue to perform difficult emotional tasks on their own. And when more forgo opportunities to perform their own emotional tasks, it becomes more likely that more would become dependent on GenAI, thereby exacerbating the problems of emotional alienation that we have outlined in the previous section.

6 Objection and Response

Before we conclude, it is worth addressing an important objection that may be raised against our argument. We have argued that outsourcing the task of emotional expression to GenAI runs the risk of cutting oneself off from the task of interpreting and expressing one's emotions for oneself, and thereby risks opting out of a crucial way of making sense of one's own values and identity. An important objection that one may raise is that our argument may prove too much and classify lots of intuitively unobjectionable ways in which people outsource the expression of emotion as forms of problematic outsourcing. For instance, we show friends that we care about them by buying them cards to congratulate them for their successes or to wish them a speedy recovery when they are sick. We might try to express the love we feel for someone by sending them a song or a poem written by someone else that we think perfectly reflects our feelings. All of these seem like unobjectionable ways in which we outsource the task of emotional expression to the objects in our environment. Yet, it seems like the argument we have given about the dangers of outsourcing emotional expression to GenAI should apply here as well. If it is dangerous to outsource emotional expression to GenAI then it seems like we should also be worried about outsourcing to cards, songs and poems written by other people. If we do not want to accept this seemingly absurd conclusion, then it seems that we should reject the claim we have defended in this paper.

There is, though, an important difference between the form of emotional outsourcing to GenAI and the forms of outsourcing considered here. As mentioned in Sect. 3, our focus is on those for whom GenAI functions as such a trusted and entrenched affective resource that it has become incorporated into their normal mode of affective agency. The kind of case we have in mind is someone who resorts to GenAI multiple times a day to outsource the task of emotional expression. This is quite different from someone who buys cards to mark special celebratory occasions like birthdays and graduations or express sympathy during periods of illness or grief.

First, there is a very clear difference in degree. Someone who occasionally draws on objects in their environment to help express their feelings is only outsourcing a

small portion of their emotional lives. They are still very much engaged in the very human process of trying to interpret and express their emotional lives and, through doing so, understand who they are more clearly. Second, there may also be a difference in kind. Someone who unthinkingly outsources an emotional expression to GenAI without checking that the expression actually matches what they are feeling is engaging in a very different kind of process from someone who painstakingly searches through all of their favourite love songs to find the one that best expresses how they feel for the person they are in love with. In this latter case, the song is carefully selected for its ability to perfectly express what one is feeling and might be thought to give a richer and more accurate representation of what the lover is feeling than their attempt to put it into words themselves could. By searching for the perfect song, they are engaging in a process of trying to interpret their own feelings and find the right way to communicate them to others. Someone who simply asks GenAI to write an apology for them and who sends it off without reading it is not engaging in this process.

What this objection and our response highlight, though, is that not all uses of GenAI for emotional expression will be ones that give rise to worries about problematic outsourcing. Those who use it sparingly and carefully will not be shutting themselves off from the project of self-interpretation. For instance, someone who asks GenAI for several different options of how to write an apology and chooses the one that best fits exactly how they are feeling may avoid some of the worries about outsourcing that we have raised here. Nevertheless, the ease, speed and ready availability of GenAI means that once we have started outsourcing our emotional expressions to it, it may become all too tempting to continue to outsource more and more of our emotional expressions. While there may be some possible uses of this technology that avoid the worries we have considered here, there remains good reason to worry about how users will actually interact with this technology.

7 Conclusion

Our emotions can reveal much to us, and our attempt to articulate our emotions can bring us to a better understanding of what we value and what we take to be core to our identity. Certain uses of GenAI – particularly, completely outsourcing the communication or articulation of our emotions – can constitute problematic outsourcing of our emotions. When we choose to outsource, we forgo important moments of self-revelation. We have further argued that the growing ready availability of GenAI creates capacity-hostile environments for our emotional self-articulation.

Against the backdrop of growing capacity-hostile environments, it is unreasonable to expect individuals to predominantly adhere to unproblematic uses of GenAI. As Ferdman (2026) rightly points out, the response to such structural issues requires structural solutions. Here, instead of attempting to take on the task of pinpointing any solution, we wish to end this paper by highlighting an important structural difficulty any solution must take into consideration.

Buijsman et al. (2025) highlight that one important safeguard against our values being led astray by GenAI is to introduce friction into AI. Instead of GenAI producing everything after the user provides the initial instructions, users' autonomy can be better preserved if the process involves GenAI constantly pausing and allowing the human user to make decisions, provide input, and take sufficient control over the generation of outputs. We believe that our core worry of problematic emotional outsourcing can be significantly alleviated should corporations introduce friction into their products when engaging in emotional tasks.

Yet, we also believe that corporations are not motivated to do so. Users generally prefer frictionless experiences (Zohar et al., 2026). It is thus quite possible that products that are more beneficial to users' emotional self-articulation would be abandoned by the majority of customers. (Consider online dating apps that require more input, and thus slower responses to one's date.) This creates individual incentives for corporations to stick to frictionless user experiences. Moreover, we believe that the commercial interests involved are massive. Corporations may be incentivised to entrap users into becoming more dependent on GenAI, to either extract more subscription fees, or to extract more personal (affective) data (Brown & Brooks, 2026). This may call for policies and regulations as solutions. However, given the borderless nature of the online world, users can always look for some other frictionless product elsewhere. We contend that any attempt to provide safeguards to users through policies and regulations must take this structural issue into consideration.

Acknowledgments We would like to thank the two reviewers for their amazing feedback that helped us bring forth a significantly better version of the paper.

Author Contribution Equal contribution.

Funding NA.

Data Availability NA.

Declarations

Ethics Approval and Consent to Participate NA.

Consent for Publication Yes.

Competing Interests NA.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Archer, A. (2024). How public statues wrong: Affective artifacts and affective injustice. *Topoi*, 43(3), 809–819.
- Bai, L., Liu, X., & Su, J. (2023). ChatGPT: The cognitive effects on learning and memory. *Brain-X*, 1(3), e30.
- Boyle, S. (2025). AI ‘wingmen’ bots to write profiles and flirt on dating apps. *The Guardian*. <https://www.theguardian.com/lifeandstyle/2025/mar/08/ai-wingmen-bots-to-write-profiles-and-flirt-on-dating-apps>
- Brewer, T. (2011). On alienated emotions. In C. Bagnoli (Ed.), *Morality and the emotions* (pp. 276–299). Oxford University Press.
- Brown, R. L., & Brooks, R. C. (2026). Smartphones: Parts of our minds? Or parasites? *Australasian Journal of Philosophy*, 104(1), 1–16.
- Buijsman, S., Carter, S. E., & Bermúdez, J. P. (2025). Autonomy by Design: Preserving Human Autonomy in AI Decision-Support. *Philosophy & Technology*, 38(3), 97.
- Campbell, S. (1997). Interpreting the personal: Expression and the formation of feelings. Cornell University Press.
- Candelon, F., Krayer, L., Rajendran, S., & Martinez, D. Z. (2023). How people can create—and destroy—value with generative AI. *BCG Glob*, 21, (2023).
- Chan, C. K. (2024). Exploring the Factors of AI Guilt Among Students—Are You Guilty of Using AI in Your Homework? ArXiv.
- Cong-Lem, N., Tran, T. N., & Nguyen, T. T. (2024). Academic integrity in the age of generative AI: Perceptions and responses of Vietnamese EFL teachers. *Teaching English with Technology*, 24(1), 28–47.
- Ferdman, A. (2026). AI deskilling is a structural problem. *AI & SOCIETY*, 41(4), 3001–3013.
- Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.
- Gallegos, F. (2022). Affective injustice and fundamental affective goods. *Journal of Social Philosophy*, 53(2), 185–201.
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6.
- Goddard, K., Roudsari, A., & Wyatt, J. C. (2011). Automation bias—a hidden issue for clinical decision support system use. *International Perspectives in Health Informatics*, 17–22.
- Goddard, K., Roudsari, A., & Wyatt, J. C. (2014). Automation bias: empirical results assessing influencing factors. *International Journal of Medical Informatics*, 83(5), 368–375.
- Hirsch, L. E. (2007). Weaponizing classical music: Crime prevention and symbolic power in the age of repetition. *Journal of Popular Music Studies*, 19(4), 342–358.
- Jakesch, M., Bhat, A., Buschek, D., Zalmanson, L., & Naaman, M. (2023). Co-writing with opinionated language models affects users’ views. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (pp.1–15).
- Kapoor, A., Diaz, M., MacNeil, S., Porter, L., & Denny, P. (2025). Exploring Student Behaviors and Motivations using AI TAs with Optional Guardrails. arXiv preprint arXiv:2504.11146.
- Kidd, C., & Birhane, A. (2023). How AI can distort human beliefs. *Science*, 380(6651), 1222–1223.
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., & Maes, P. (2025). Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task. arXiv preprint arXiv:2506.08872.
- Krueger, J. (2023). An ecological approach to affective injustice. *Philosophical Topics*, 51(1), 85–112.
- Lai, T. H. (2025). Racist Monuments: The Beauty is the Beast. *The Journal of Ethics*, 29(1), 21–41.
- Lin, T. A., & Chen, P. H. C. (2022). Artificial intelligence in a structurally unjust society. *Feminist Philosophy Quarterly*, 8(3/4).
- Lin, T. A., & Huang, L. T. L. (2026). AI, Normality, and Oppressive Things. *Minds and Machines*, 36(2), 26.
- Osler, L., Engelen, B., & Archer, A. (2025). The ethics and politics of nudges and niches: A critical analysis of exclusionary environmental designs. *Crime prevention by exclusion* (pp. 182–206). Routledge.
- Pallie (n.d.) Breakup Text Generator. Available at: <https://relationship.pallie.ai/free-tools/breakup-text-generator/>. Accessed 26 Aug 2026
- Parker, T. (Ed.). (2023). (Director) Deep Learning. South Park [TV series] Season 26 Episode 4 (3/8/2023) South Park Studios.

- Piredda, G. (2020). What is an affective artifact? A further development in situated affectivity. *Phenomenology and the Cognitive Sciences*, 19(3), 549–567.
- Pitts, G., Marcus, V., & Motamedi, S. (2025). Student perspectives on the benefits and risks of AI in education. arXiv. arXiv preprint arXiv:2505.02198.
- Srinivasan, A. (2018). The aptness of anger. *Journal of Political Philosophy*, 26(2), 123–144.
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386.
- Tait, A. (2024). My boyfriend sent me a heartfelt message - then I learned he'd used ChatGPT. The iPaper. <https://inews.co.uk/inews-lifestyle/boyfriend-sent-heartfelt-message-learned-chatgpt-3413336>
- Timms, R., & Spurrett, D. (2023). *Hostile scaffolding Philosophical Papers* 52(1), 53–82.
- Tobi, A. (2025). The Epistemic Harms of Botched Apologies for Past Wrongs. *Journal of Applied Philosophy*.
- Vicente, L., & Matute, H. (2023). Humans inherit artificial intelligence biases. *Scientific Reports*, 13(1), 15737.
- Weirich, K., & Holdier, A. G. (2026). Emotional outsourcing (p. pqag043). *The Philosophical Quarterly*.
- Wessel, N. C. (2023). Decision-support systems and decision making: Managing decisional deskilling in human-DSS interactions in organizations. In ICDS 2023: The Seventeenth International Conference on Digital Society.
- Williams-Ceci, S., Jakesch, M., Bhat, A., Kadoma, K., Zalmanson, L., & Naaman, M. (2026). Biased AI writing assistants shift users' attitudes on societal issues. *Science Advances*, 12(11), eadw5578.
- Zhang, L., & Xu, J. (2025). The paradox of self-efficacy and technological dependence: Unraveling generative AI's impact on university students' task completion. *The Internet and Higher Education*, 65, 100978.
- Zohar, E., Bloom, P., & Inzlicht, M. (2026). Against frictionless AI. *Communications Psychology*, 4(1), 39.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.