Everyday moral transgressions (EMTs): Investigating the morality of everyday behaviors

J. Dean Elmore[1] · Jerome A. Lewis[2] · Zachary M. Himmelberger[3] · Jefferson A. Sherwood[4]

Abstract

The vignette is a mainstay of contemporary psychological research on morality. However, many commonly used moral vignettes depict situations that people are unlikely to encounter in their daily lives. In this exploratory investigation, we sought to develop a series of moral vignettes that more closely resemble everyday life. In Study 1, our aim was to assemble an inventory of common situations that arouse people's moral concerns. Participants read 70 vignettes and indicated whether the behaviors depicted therein were morally relevant. In Study 2, we compared the "most immoral" vignettes from Study 1 to a series of vignettes from the moral psychology literature. As expected, participants rated the behaviors depicted in our vignettes as being less morally wrong but more typical than those depicted in existing stimuli. These findings indicate that many everyday events arouse people's moral concerns and suggest that stimuli like our everyday moral transgressions may be of great utility to researchers studying the psychology of morality.

Keywords Morality · Moral psychology · Moral judgment · Vignette

Although morality has been a topic of interest for psychologists for some time, there has been renewed interest in this topic over the past few decades in response to new theoretical and empirical developments. This renewed interest has led to a dramatic increase in the number of studies concerning moral psychology and related topics. This growing body of research has been multi-faceted, with several lines of research attempting to address different aspects of morality, including moral reasoning, moral emotions, and moral judgment (see Ellemers et al., 2019, for a review). Although researchers have utilized various methodological tools to study the psychology of morality, one common approach to testing these types of questions is to present participants with a moral scenario (in the form of a vignette) and then assess their responses. This approach has proven exceedingly fruitful in advancing our understanding of the various psychological and behavioral processes related to morality. Yet, despite the widespread use of vignette studies in moral psychology, existing experimental stimuli have key limitations that make them less effective for addressing certain types of research questions. Specifically, many existing vignettes depict situations that are unusual or uncommon. In this exploratory investigation, we sought to address this limitation by developing a set of novel vignettes that are more representative of people's lived experiences related to morality.

Even as psychologists have begun to ask increasingly diverse and more nuanced questions about morality, the general format of the vignettes that psychologists employ to test these questions has remained relatively consistent. Most moral vignettes take the form of brief scenarios in which third-party individuals perform behaviors that carry some degree of moral weight. In some cases, researchers have devised novel moral vignettes to achieve the goals of a specific project (e.g., Rottman et al., 2014). However, in many cases, researchers have looked to Haidt and colleagues' work on moral foundations theory (MFT; e.g., Haidt, 2001; Haidt, 2007; Graham et al., 2011) as a source of inspiration

✉ J. Dean Elmore
jelmore@benedictine.edu

1   Department of Psychological Sciences, Benedictine College, 1020 N 2nd St, Atchison, KS 66002, USA

2   Department of Psychology, Bellevue University, Bellevue, NE, USA

3   Maryville College, Division of Behavioral Sciences, Maryville, TN, USA

4   School of Psychology, University of Southern Mississippi, Hattiesburg, MS, USA

for their vignettes. Vignettes may serve different purposes across studies, with some researchers using vignettes to deliver independent variable manipulations (e.g., Hirozawa et al., 2020) and others employing vignettes to capture dependent variable responses (e.g., Eskine et al., 2011). Thus, moral vignettes (particularly those based on the MFT) have been used to answer a wide range of questions related to morality, such as whether the physical experience of disgust affects moral judgment (Wheatley & Haidt, 2005), whether moral/immoral behavior influences perceptions of physical attractiveness (He et al., 2022), and what factors affect decisions about who to sacrifice in moral dilemmas (Białek et al., 2018), among many others (e.g., Dong et al., 2023; Isler et al., 2021; Ochoa et al., 2022; Parkinson et al., 2011; Pennycook et al., 2014; Russell and Giner-Sorolla, 2011; Salerno & Peter-Hagene, 2013; Sunar et al., 2021).

A recurring concern with existing moral vignettes is that they depict scenarios that participants are unlikely to encounter in day-to-day life. This is particularly true for vignettes based on the purity foundation described in MFT (Graham et al., 2009). For instance, one commonly used purity vignette describes a man cooking and eating his dog after it died of natural causes. In another purity vignette, an individual is described as undergoing plastic surgery in order to add a 2-inch tail onto the end of their spine. These are highly unusual events that participants likely have no previous experience with. Indeed, research by Gray and Keeney (2015) suggests that people view scenarios like these as significantly "weirder" than other, more "naturalistic" purity-based scenarios (e.g., someone having an affair while still married to someone else).

A case could also be made that many of the commonly used moral vignettes, such as those based on the harm foundation of MFT, are unrepresentative of daily life because they depict events that are rare or severe. For instance, one harm scenario describes someone shooting and killing a member of an endangered species. In another harm scenario, a person is described as sticking a pin into the palm of a child they do not know (Haidt, 2007). Although these events are less stereotypically weird than the purity scenarios and are in a strict sense plausible, they are not (at least in our view) events that people are likely to encounter in everyday life, which is when many of our moral experiences occur (e.g., Hofmann et al., 2014).

To be clear, we are not arguing against the use of moral vignettes that depict atypical acts or ones that entail more severe outcomes. In many cases, these types of vignettes may be well-suited to the research question that is being addressed. For example, Parkinson et al. (2011) examined whether moral judgments related to harm, dishonesty, and disgust produce a unified response in the brain or a more differentiated one. They presented participants with a variety

of moral vignettes related to the three areas of concern. Although some of the vignettes depicted situations that are more commonplace (e.g., a parent reprimanding their children with a ruler), others were less representative of everyday life (e.g., someone engaging in sexual intercourse with a dead animal). For Parkinson et al. (2011), the use of extreme or unusual situations seems justified and indeed may have been necessary given that the goal of the research was to test whether different moral judgments produce distinct patterns of neural activation. If they had exclusively used scenarios that more closely mirrored everyday life, then the manipulation might not have been strong enough to produce the observed results. However, as previously noted, these types of scenarios might not be appropriate for all research questions, such as the role that daily interactions play in people's judgments about the moral character of their neighbors and colleagues.

The widespread use of vignettes in moral psychology has led some researchers to create more standardized sets of moral vignettes (e.g., Knutson et al., 2010). One standout resource comes from Clifford et al. (2015), who provide a rigorously tested and robust set of vignettes that can be used for an array of research purposes (though they placed particular emphasis on creating vignettes that were suitable for neuroimaging studies). The 90 vignettes that they developed have several benefits. First, their scenarios map neatly onto the moral concerns identified in the MFT (e.g., care, fairness, loyalty; Graham et al., 2011). Second, their scenarios are numerous, diverse, and largely free from cultural bias. Finally, their scenarios are uniform in their format, with most scenarios containing 60 to 70 characters, and were tested to ensure their readability and comprehensibility (Clifford et al., 2015). Importantly, many of their scenarios depict events that are less outlandish and less severe than those mentioned previously. Thus, Clifford et al.'s (2015) moral foundations vignettes (MFVs) improve on existing experimental stimuli in several ways.

Despite these improvements, the MFVs have key drawbacks. First, many of the vignettes developed by Clifford et al. (2015) specify an agent who performs a moral act and a patient who experiences the effects of that act. Although this characteristic of the MFVs comports with theories suggesting that the agent-patient relationship is essential for an event to be considered moral in nature (i.e., the theory of dyadic morality; Gray et al., 2012; Schein & Gray, 2018), other theories do not make such a stipulation, instead suggesting people may learn to categorize various things (e.g., events, behaviors, etc.) as "morally right" or "morally wrong" over time and in relation to different contextual cues (McHugh et al., 2022). Thus, it is possible that the MFVs miss a wide range of behaviors and events for which there is

no salient moral dyad, but that nevertheless may be considered moral in nature.

The second concern is that across the MFVs subscales, participants rated the vignettes as being low on "frequency," indicating that they did not frequently encounter the depicted actions in daily life (Clifford et al., 2015, study 1). It is important to acknowledge that Clifford et al. (2015) did not explicitly set out to produce vignettes that depict routine behaviors that might be considered morally relevant. Rather, they intended to produce a set of vignettes that were consistent in their structure, that mapped well onto the moral foundations, and that could easily be used in neuroimaging studies. In those aims, they seem to have succeeded. That said, the low frequency ratings raise concerns about the extent to which the MFVs can be used to study the ways that people make moral judgments in their daily lives.

The current research

The vignette is a mainstay of modern moral psychology. However, as we have outlined, existing experimental stimuli have characteristics that may make them unsuitable for certain questions related to morality. In particular, many of the most commonly used moral vignettes depict behaviors or events that could be considered (1) atypical or "weird" (Clifford et al., 2015, study 1; Gray & Keeney, 2015), (2) that are exceptionally rare or severe, or (3) that have a clear moral agent and a clear moral patient (i.e., a victim). Although such scenarios are possible, we frequently encounter behaviors and events that do not have a clear victim or whose consequences may be more benign. Thus, many of the vignettes used in previous research fail to depict the sorts of moral situations that people find themselves in over the course of everyday life.

In this investigation, we sought to expand on previous research by evaluating participants' reactions to a variety of moral scenarios that we believe more closely resemble people's daily experiences. We conducted two studies in pursuit of this goal. Because most studies only evaluate people's reactions to a handful of scenarios, in Study 1, we sought to assess the perceived morality (or immorality) of a wide range of everyday situations. In Study 2, we took the 15 scenarios that were most often identified as being "immoral" and compared their severity and typicality to some of the existing moral vignettes identified above (i.e., Clifford et al., 2015; Graham et al., 2009).

Study 1

The primary aim of Study 1 was to determine whether various routine behaviors and events encountered in everyday life activate people's sense of morality and, if so, which of these behaviors and events would receive the strongest endorsement as having moral weight. As other researchers have done (e.g., Clifford et al., 2015), we began by writing as many scenarios as possible. Throughout this iterative process, we reviewed each proposed scenario as a group and dismissed or reworked those that were ambiguous or unclear (e.g., due to wording). In keeping with our broader research goals, we first and foremost attempted to write scenarios that were representative of day-to-day life. We drew inspiration for the vignettes from personal experience, from existing moral vignettes, from well-known injunctive norms (e.g., Jacobson et al., 2011), and from other contemporary cultural depictions of immoral or rude behavior (e.g., failing to put a shopping cart back; Hauser, 2021). When possible, we ensured that the focal behavior depicted in the vignette caused minimal to no direct harm, had no clear victim, and could easily be performed by a wide range of actors (i.e., were independent of various demographic characteristics). In scenarios in which a harmed party could be inferred, the nature of the harm was not severe, more closely resembling minor annoyance or inconvenience. We initially produced 72 scenarios that met these general criteria. These scenarios were then reviewed by six undergraduate research assistants at an institution that did not provide participants for either study. The research assistants' primary task was to indicate whether they believed each scenario had the potential to be viewed as morally relevant by a wider audience and whether each scenario could believably occur in everyday life. In addition to these judgments, the research assistants also provided informal feedback about the clarity and wording of our scenarios. Based on their responses, we removed two scenarios, leaving 70 scenarios for participants to evaluate in Study 1.[1]

Participants

Our sample comprised students from Bellevue University ($N = 79$). Demographic information for the sample was collected in a separate pre-test survey. Because some participants failed to provide the information necessary to link their responses on the pre-test survey to their responses in

---

[1] Aside from our own internal review process and the feedback from the undergraduate research assistants, we did not subject our vignettes to any tests of readability or comprehensibility prior to conducting Study 1.

Study 1,[2] we only retained demographic information for 32 participants, though all 79 participants evaluated the moral scenarios. Most of the participants for whom we obtained demographic information were women (59.30%) and were White/non-Hispanic (56.25%). The mean age of these participants was 35.41 ($SD = 9.28$, min $= 23$, max $= 52$). Of those same participants, 37.5% reported being affiliated with the Democratic party, 25% reported being affiliated with the Republican party, 15.6% identified as Independent, and 21.9% reported another political affiliation. All participants were recruited through undergraduate courses and were given extra credit or course credit for completing the study.

Materials and Procedure

Before participating in Study 1, participants completed a pre-test survey, including a demographics questionnaire. Students who completed the pre-test survey were eligible to participate in Study 1. Participants were first asked to read an informed consent form. After agreeing to proceed with the study, participants were told that they would be reading and responding to scenarios depicting common behaviors. Aside from one scenario involving a romantic couple, all scenarios involved a single person engaged in a discrete instance of behavior. Participants were shown 70 scenarios in total, and the order of vignettes was randomly determined for each participant. For each scenario, participants were asked to indicate whether the events depicted were "moral," "immoral," or "not related to morality." This rating task was kept deliberately simple for two reasons. First, whereas researchers in previous studies have asked participants to rate a subset of their total number of vignettes (e.g., 14 to 16 vignettes out of 132 total; Clifford et al., 2015, study 1), we asked participants to rate all 70 of the vignettes that we created. If we had asked for more complex ratings or had included more ratings alongside each scenario (e.g., judging them according to the moral foundations), then participants might have been overburdened by the study protocols. So, in part, we adopted a simple response measure to avoid participant fatigue. Second, because the goal of this study was to evaluate the moral relevance of a range of everyday situations (as opposed to the degree of their moral wrongness), we reasoned that forcing a categorical response would allow us to more readily obtain a consensus about the morality of each situation. A full list of the scenarios and their Study 1 ratings can be found in the Supplementary Material. After viewing all the scenarios, participants were debriefed as to the purpose of the research. These protocols were approved by the local institutional review board (IRB).

Results and discussion

The data and R code for Study 1 can be found at https://github.com/ZacharyHimmelberger/everyday-moral-transgressions. We determined the moral relevance of each scenario by assessing the percentage of participants who labeled the scenario as "moral," "immoral," or "not related to morality." Participants displayed considerable variability in their tendency to view the depicted behaviors in moral terms. Indeed, whereas some participants endorsed as few as zero scenarios as carrying moral weight (i.e., as being "moral" or "immoral"), others viewed as many as 69 scenarios as being morally relevant.[3] Despite this variability among participants, several of the behaviors depicted in the scenarios were widely viewed as being immoral. For instance, 93.15% of participants reported that it was immoral to ignore an approaching ambulance and to continue driving without slowing down or getting out of the way. Similarly, 89.19% of participants indicated that it was immoral to urinate on a toilet seat in a public restroom and then leave without cleaning it up. Table 1 displays the 15 scenarios that received the strongest endorsement as being immoral.

Although we were primarily interested in determining which behaviors would be viewed as "immoral," we also considered which scenarios participants viewed as being "moral." Of the 70 scenarios, only 14 scenarios were endorsed as being moral at a rate of 10% or higher. The only vignette to receive greater than 30% endorsement as moral depicted a scenario in which someone takes off work for a mental health day (68.06% endorsed as "moral"). After that scenario, the two most "moral" scenarios depicted someone keeping a ten-dollar bill that they found outside their apartment building (29.73%) and someone drinking beer in a public place (22.22%).

Lastly, we evaluated which scenarios participants were most likely to view as "not related to morality." The scenarios that most frequently fell into this category included someone playing loud music at a stoplight (73.97%), someone sending a message to a colleague that contains numerous typographical errors (69.44%), and someone using the restroom in a local retail shop and then leaving without making a purchase (66.67%).

The findings from Study 1 indicate that people view many day-to-day behaviors as being morally relevant. This was especially true for several scenarios that were strongly

---

[2] Specifically, participants were given a research ID at the end of the pre-test survey that they were asked to provide in subsequent studies. As noted, many participants failed to provide their research ID, instead providing other sorts of information (e.g., student ID, name).

[3] Exploratory analyses revealed significant associations between these individual differences in moral judgment and other meaningful variables (e.g., age). See Supplementary Material for the full results.

endorsed as being "immoral." Further, we found that some behaviors were widely viewed as being either "moral" or "unrelated to morality."

Study 2

In Study 2, we compared the vignettes that we developed for Study 1 to various vignettes used in previous research. In particular, we compared our vignettes to those developed by Clifford et al. (2015) and to those based on the MFT (Graham et al., 2009, study 3). As previously noted, scenarios based on the MFT have been rated by non-researchers as being "weirder" than other more "naturalistic" scenarios (Gray & Keeney, 2015). Similarly, though many of Clifford et al.'s (2015) MFVs depict less outlandish scenarios, on average, participants rated the scenarios as occurring infrequently. Thus, our expectation in Study 2 was that participants would view the scenarios from Study 1 as being more typical or commonplace than the existing vignettes. We also anticipated that because they depict more routine behaviors, the scenarios from Study 1 might be viewed as less severe or as less morally wrong.

Participants

We initially received 285 responses to the Study 2 survey. However, 15 participants were removed for excessive missing data (i.e., 20 or more items skipped), for completing the survey more than once, or upon request following debriefing. Our final sample ($N = 270$) comprised students from Benedictine College ($n = 197$) and Bellevue University ($n = 73$).[4] Participants were 21.89 years old on average ($SD = 7.01$, min $= 17$, max $= 58$). Most participants described themselves as women (71.9%), and most participants were White/non-Hispanic (72.59%). Sixty-two percent of participants reported being affiliated with the Republican party, 17.3% reported being affiliated with the Democratic party, 14.7% identified as Independent, and 6.0% reported another political affiliation. Participants were recruited through their undergraduate classes and were offered extra credit or course credit for their participation. To ensure that all students had equal opportunity to participate, we did not place constraints on sample size, instead accepting responses passively for a pre-determined time period. As a result, we did not conduct a priori power analyses. However, a post hoc sensitivity analysis indicated that we had sufficient power to detect small effects ($f = 0.07$) in our most complex analysis (i.e., a repeated-measures ANOVA with seven measurements; Faul et al., 2007).

---

[4] Bellevue students who participated in Study 1 were not permitted to participate in Study 2.

Materials and Procedure

Participants navigated to the study website using a link sent to them by professors at the participating institutions. Participants first read an informed consent form followed by an overview of the study and a set of general instructions. Participants were told that the purpose of the study was to assess people's general reactions to various moral vignettes that have been used in previous research. The instructions indicated that participants would be asked to read 71 vignettes and rate each on two dimensions: (1) the moral wrongness of the behaviors depicted in the vignette, and (2) how likely they would be to encounter the depicted behavior in day-to-day life (henceforth referred to as "typicality"). Ratings on both measures were made on a Likert-type scale ranging from 0 ("not at all") to 5 ("extremely").

Next, participants read and rated the vignettes. To obtain an adequate comparison of the vignette groups in terms of their typicality, we reasoned that it was necessary to match the vignettes as closely as possible in terms of their moral valence. Because many of the existing vignettes have previously been rated as morally wrong (i.e., Clifford et al., 2015, study 1), we decided to only include those of our vignettes that were widely viewed as being immoral in Study 1 (i.e., receiving greater than 75% endorsement). Fifteen of our vignettes met that criterion and did not overlap substantially with other scenarios that were strongly endorsed as being immoral in Study 1. For the remainder of the paper, we will refer to these 15 vignettes as the everyday moral transgressions (EMTs). The text of these vignettes and their ratings from Study 2 (i.e., means and *SD*s) appear in Table 1. We also included vignettes based on the harm, fairness, and purity taboo trade-off items from study 3 of Graham et al. (2009). The text of these scenarios was modified slightly so that each scenario described an unidentified person performing an act (e.g., "Someone kicks a dog in the head, hard"). Lastly, we included the emotional harm, physical harm, fairness, and purity vignettes produced by Clifford et al. (2015). These were not placed in a common format with the other sets of vignettes but were instead left in their original format. Thus, these vignettes retained references to specific actors (e.g., a girl laughing at her friend's dad because of his occupation). Vignettes were presented in a randomized order for each participant. After reading and rating all vignettes, participants completed a demographics questionnaire. These protocols received approval from the IRBs at both participating institutions.

Results and discussion

We conducted one-way repeated measures ANOVAs using the *afex* package in R (Singmann et al., 2023) to compare

Table 1 Mean (SD) Moral Wrongness and Typicality Ratings for the "Most Immoral" Scenarios (As Indicated by Study 1 Endorsement)

Vignette	Study 1	Study 2	
	% rating the scenario as immoral	Mean (SD) wrongness rating	Mean (SD) typicality rating
1. Someone ignores an approaching ambulance and continues driving without slowing down or getting out of the way.	93.15	3.90 (1.10)	2.14 (1.47)
2. Someone cheats on a college exam.	93.15	3.66 (1.17)	2.70 (1.60)
3. Someone urinates on the toilet seat in a public restroom and leaves without cleaning it up.	89.19	2.65 (1.43)	2.95 (1.63)
4. Someone silently watches pornography at their cubicle during a lull at work.	87.84	4.19 (1.15)	1.43 (1.44)
5. Someone sneaks into the middle of a long line.	85.14	2.32 (1.32)	3.26 (1.55)
6. Someone answers their cell phone and has a conversation in a movie theater.	83.78	2.03 (1.30)	1.87 (1.17)
7. Someone allows elevator doors to close on another person who is hurrying to catch it.	82.19	2.16 (1.37)	2.19 (1.33)
8. Someone throws a paper cup on the ground instead of taking it to a trash can.	81.08	2.58 (1.35)	3.15 (1.60)
9. Someone's dog poops in a public park and they leave without picking it up.	80.82	2.37 (1.33)	3.08 (1.58)
10. Someone ignores a sign asking pedestrians to avoid walking on the grass.	80.56	1.98 (1.33)	2.79 (1.58)
11. Someone takes up more space than they really need on a subway car that is full of people.	80.56	2.30 (1.31)	2.34 (1.58)
12. Someone sees a colleague drop a large stack of papers and quickly walks in the opposite direction before they are seen.	80.28	2.33 (1.33)	2.09 (1.45)
13. Someone leaves their grocery cart in their parking space instead of returning it to the cart return.	79.73	2.05 (1.36)	3.57 (1.58)
14. Someone drops a container of eggs on the floor in a grocery store and then puts it back on the shelf before continuing shopping.	79.73	2.80 (1.27)	2.16 (1.47)
15. Someone drives behind another motorist with their high beams on.	78.38	2.50 (1.46)	2.86 (1.57)

Table 2 Descriptive Statistics for All Vignette Groupings

Vignettes	Moral wrongness		Typicality	
	M (SD)	95% CI	M (SD)	95% CI
EMT items	2.68 (0.95)	2.56–2.79	2.59 (1.06)	2.47–2.72
Clifford et al. (2015) - total	3.37 (0.79)	3.27–3.46	1.52 (0.85)	1.42–1.62
Harm vignettes	3.01 (1.03)	2.88–3.13	1.74 (0.98)	1.62–1.86
Fairness vignettes	3.27 (0.85)	3.17–3.37	1.96 (0.99)	1.85–2.08
Purity vignettes	4.15 (0.76)	4.06–4.25	0.60 (0.72)	0.52–0.69
Graham et al. (2009) - total	3.34 (0.69)	3.26–3.43	1.25 (0.77)	1.15–1.34
Harm vignettes	3.47 (0.83)	3.37–3.57	1.63 (0.92)	1.52–1.74
Fairness vignettes	3.84 (0.71)	3.75–3.93	1.58 (1.03)	1.45–1.70
Purity vignettes	2.72 (1.08)	2.59–2.85	0.52 (0.66)	0.44–0.60

judgments of wrongness and typicality across the three vignette groups, namely, those from the current study (the EMTs), those from Clifford et al. (2015), and those based on Graham et al. (2009, study 3). To conduct these analyses, we calculated a mean score for each vignette group for each participant. This was done for both moral wrongness and typicality ratings. If a participant was missing one or more items from one of the vignette groups, we calculated their score(s) based on the items for which they provided responses.[5] The data

and R code for these analyses can be found at https://github.com/ZacharyHimmelberger/everyday-moral-transgressions. The means and standard deviations for these vignette groups are presented in Table 2. In the comparison of moral wrongness, Mauchly's test indicated a violation to the assumption of sphericity, $\chi^2 (2) = 0.81$, $p < .001$, so degrees of freedom were corrected using Greenhouse-Geisser ($\epsilon = .84$). There was a significant main effect of vignette

---

[5] Most of our participants had no missing values (n = 204) or were only missing two or fewer values (n = 53). Thus, most participants

(95%) had complete or nearly complete data. Only 13 participants had missing data for three or more items (up to the missing data cutoff of 20). Of those 13 participants, only two skipped more than eight items.

group on wrongness, $F(1.68, 451.10) = 350.58$, $p < .001$, $\eta_p^2 = .57$.[6] Post hoc tests using a Bonferroni correction revealed that, on average, the EMTs were judged as less morally wrong than the items from Graham et al. and Clifford et al. (ps < .001). There was no significant difference between the latter two vignette groups ($p = .35$). In the comparison of typicality, Mauchly's test again indicated a violation, $\chi_c^2(2) = 0.52$, $p < .001$, so degrees of freedom were corrected using Greenhouse-Geisser ($\epsilon = .68$). There was a significant main effect of vignette group on typicality, $F(1.35, 363.05) = 785.86$, $p < .001$, $\eta_p^2 = .75$. Post hoc tests using a Bonferroni correction revealed that all three vignette groups were significantly different from one another such that, on average, the EMTs were judged as most typical and the items from Graham et al. as least typical (ps < .001).

Each of the items from Graham et al. (2009) and Clifford et al. (2015) was categorized as relating to either harm, fairness, or purity. For thoroughness, we repeated the above analyses to compare judgments of wrongness and typicality across each of the three Graham et al. subsets, the Clifford et al. subsets, and our EMT items. The results were largely consistent with our above findings, so we do not report the ANOVAs here. Bonferroni corrected post hoc testing revealed that our EMT items were judged as significantly less wrong than all subsets (ps < .001) except the Graham et al. purity items ($p = .47$) and as significantly more typical than all subsets (ps < .001). The full details of these additional analyses can be found in the Supplementary Material.[7]

The results of Study 2 suggest that there is a push and pull between moral wrongness and typicality. Whereas the EMTs were viewed as being more typical of participants' day-to-day experiences, they were also viewed as being less morally wrong overall.

General discussion

In the current investigation, we explored whether people assign moral weight to everyday scenarios and how these evaluations compare to evaluations of moral scenarios commonly used in moral psychology research. The results of Study 1 indicated that people consistently judge many everyday scenarios as having moral weight. The results

of Study 2 showed that participants judged our everyday moral scenarios to be both more common (e.g., more typical) and less heinous than the commonly used sets of moral scenarios. Together, these findings underscore the value of considering everyday moral transgressions in research on moral judgment and provide evidence of the limited representativeness of existing moral vignettes.

In addition to emphasizing the importance of everyday behaviors to moral judgment, our scenarios could serve as stimuli for subsequent research on morality. In particular, our findings may be useful in that they provide insight into which everyday behaviors are most universally identified as being "immoral" (or "moral"; Study 1) and which are regarded as being most morally wrong (Study 2). Moreover, our scenarios may be valuable to researchers because of their flexibility. Whereas other moral vignettes often specify the characteristics of the actors in the vignette (e.g., a boy insulting a woman's appearance; Clifford et al., 2015), our vignettes do not specify the identity or characteristics of the actors, thus allowing researchers to modify them to meet their specific research needs. For instance, one of the scenarios that we asked participants to evaluate involved an individual silently watching pornography at their desk during a lull at work. Although this behavior was widely viewed as being "immoral," and received high moral wrongness ratings in the current investigation, how people perceive and respond to the depicted behavior may ultimately depend on the characteristics of the individual performing it, such as whether the actor is a man or a woman (e.g., Hester & Gray, 2020). Using our stimuli, individual researchers could easily modify the scenario to test these sorts of questions, with the ratings we have obtained in these studies serving as a reasonable baseline for assessing the scenario's moral wrongness in general.

Additionally, to the extent that harm is necessary for an act to be viewed as immoral, and a moral patient is necessary for harm to be perceived (Schein & Gray, 2018), our scenarios allow the researcher to specify a patient (vs. one being pre-determined by the text of the scenario). Compared to the more commonly used scenarios in which the characteristics of the actor(s) and patient(s) cannot be removed without losing the meaning of the scenario (e.g., a zoo trainer jabbing a dolphin to get it to perform; Clifford et al., 2015), our scenarios can be modified to explore the influence of specific agents and patients on moral judgments. In other words, our scenarios provide a base vignette than can be easily manipulated to explore the components of an actor/patient moral dyad without confounding the original context of the scenario.

It is worth noting that the scenarios we included in the current investigation are not entirely novel in the sense that some of the moral events and behaviors that we described

---

[6] Preliminary analyses indicated that this effect was moderated by institution. Because the pattern of means was similar across institutions, and there was no moderation for typicality, we collapsed across samples in the main analyses. See Supplementary Material for full details.

[7] We conducted a separate set of analyses using only the data provided by non-traditional students ($n = 48$) whose mean age was 34.35 (median = 32.50, min = 23, max = 58). The results of those analyses were similar to those of the full sample, both for the main analyses and the subset comparisons.

appear elsewhere in the literature. However, in most of these cases, everyday situations like the ones depicted in our vignettes have been used to test some other research question (e.g., whether hypnotic disgust influences moral judgment; Wheatley & Haidt, 2005). To our knowledge, no studies have attempted to evaluate such a wide range of everyday situations both in terms of their moral wrongness and their typicality. Further, when typicality has been considered in the development of moral vignettes, it has not been a primary focus but rather one of many concerns, such as the uniformity of the vignettes or their depiction of acts from different moral domains (e.g., harm vs. purity; Clifford et al., 2015).

Although the scenarios we created seem to both tap into people's daily experiences and activate their moral concerns, we did not attempt to classify our scenarios according to the different moral foundations (e.g., harm/care, purity/sanctity; Graham et al., 2011), which could possibly limit their utility for researchers who are interested in asking domain-specific questions. We adopted this approach for two reasons. First, given the goals of our research, we prioritized capturing moral events that people regularly encounter in their daily lives over ensuring the validity of our scenarios with regard to MFT. By keeping the research protocols deliberately simple and minimizing the number of ratings that participants had to make, we hoped to maintain that more narrow focus. Second, many of the existing moral psychology vignettes are already linked to MFT. For instance, Clifford et al. (2015) went to great lengths to ensure that each of their scenarios tapped into specific moral concerns and not others. Thus, the strength of our scenarios and their unique contribution to the literature stems from their representativeness regarding daily life.

As stated above, we believe that our findings raise important questions about the ecological validity of moral vignettes, and we hope that our scenarios may be of use to other researchers. However, we want to emphasize that this investigation was exploratory in nature and that additional work needs to be done for our scenarios to be considered a robust alternative to existing stimuli. Indeed, the current research is limited in a number of ways. First, the generalizability of our findings may be limited due to our use of relatively small convenience samples in both studies. Although our samples were reasonably diverse in terms of age, all participants were students from academic institutions in the Midwestern United States. In future studies, our scenarios should be evaluated by larger and more diverse samples. Second, due to the direct nature of our response measures, we have a relatively narrow perspective on participants' reactions to our scenarios. There are potentially many other questions that could be asked about the situations that we

included in our studies (e.g., numerical estimates of how frequently the depicted behaviors are encountered).

Another limitation is that many of our vignettes are culturally specific. Although moral psychology research has found some cross-cultural similarities in moral evaluations (e.g., Doğruyol et al., 2019), there are other studies indicating important cross-cultural differences in moral reasoning and behavior (e.g., Lo et al., 2020; Rhim et al., 2020), which limits our ability to evaluate the usefulness of our scenarios for research conducted in non-Western countries. A final limitation of our study is that by not identifying specific actors in the vignettes, we allowed participants to fill in the blanks, which they might have done in biased ways (e.g., by assuming that the person watching pornography at their desk is male). Although this introduces unsystematic variability into our results, it does not seem to have negated the effects that we observed in Study 2.

People have rich and varied moral experiences. On the one hand, people view many exceedingly immoral behaviors and events as rare. A large body of literature indicates that these phenomena stoke people's moral emotions and influence their moral reasoning. On the other hand, people view everyday life as full of minor, nontrivial moral phenomena. There is far less research on this. Given that such phenomena comprise a meaningful part of the typical individual's daily life, there is a clear need to study them. The current investigation emphasizes this point and represents a first step toward developing a tool with which to evaluate everyday moral experiences more closely.

Author contributions J. Dean Elmore – Conceptualization; Investigation; Methodology; Project Administration; Resources; Writing – Original Draft Preparation; Writing – Review & Editing. **Jerome A. Lewis** – Conceptualization; Investigation; Resources; Writing – Original Draft Preparation; Writing – Review & Editing. **Zachary M. Himmelberger** – Data Curation; Formal Analysis; Methodology; Validation; Writing – Original Draft Preparation; Writing – Review & Editing. **Jefferson A. Sherwood** – Resources; Writing – Review & Editing.

Declarations

References

Białek, M., Fugelsang, J., & Friedman, O. (2018). Choosing victims: Human fungibility in moral-decision making. *Judgment and Decision Making*, *13*(5), 451–457. https://doi.org/10.1017/S193029750000872X.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*, *47*(4), 1178–1198. https://doi.org/10.3758/s13428-014-0551-2.

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences*, *151*, 109547. https://doi.org/10.1016/j.paid.2019.109547.

Dong, M., Kupfer, T. R., Yuan, S., & van Prooijen, J. (2023). Being good to look good: Self-reported moral character predicts moral double standards among reputation-seeking individuals. *British Journal of Psychology*, *114*(1), 244–261. https://doi.org/10.1111/bjop.12608.

Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*, *23*(4), 332–366. https://doi.org/10.1177/1088868318811759.

Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science*, *22*(3), 295–299. https://doi.org/10.1177/0956797611398497.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029. https://doi.org/10.1037/a0015141.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366. https://doi.org/10.1037/a0021847.

Gray, K., & Keeney, J. E. (2015). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*, *6*(8), 859–868. https://doi.org/10.1177/1948550615592241.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*(2), 101–124. https://doi.org/10.1080/1047840X.2012.651387.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*(5827), 998–1002. https://doi.org/10.1126/science.1137651.

Hauser, C. (2021). June 8). Everyone has a theory about shopping carts. *The New York Times*. https://www.nytimes.com/2021/06/08/style/shopping-cart-parking-lot.html.

He, D., Workman, C. I., He, X., & Chatterjee, A. (2022). What is good is beautiful (and what isn't, isn't): How moral character affects perceived facial attractiveness. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000454.

Hester, N., & Gray, K. (2020). The moral psychology of raceless, genderless strangers. *Perspectives on Psychological Science*, *15*(2), 216–230. https://doi.org/10.1177/1745691619885840.

Hirozawa, P. Y., Karasawa, M., & Matsuo, A. (2020). Intention matters to make you (im)moral: Positive-negative asymmetry in moral character evaluations. *The Journal of Social Psychology*, *160*(4), 401–415. https://doi.org/10.1080/00224545.2019.1653254.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*, 1340–1343. https://doi.org/10.1126/science.1251560.

Isler, O., Yilmaz, O., & Doğruyol, B. (2021). Are we at all liberal at heart? High-powered tests find no effect of intuitive thinking on moral foundations. *Journal of Experimental Social Psychology*, *92*, 104050. https://doi.org/10.1016/j.jesp.2020.104050.

Jacobson, R. P., Mortensen, C. R., & Cialdini, R. B. (2011). Bodies obliged and unbound: Differentiated response tendencies for injunctive and descriptive social norms. *Journal of Personality and Social Psychology*, *100*(3), 433. https://doi.org/10.1037/a0021470.

Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., Adolphs, R., & Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience*, *5*(4), 378–384. https://doi.org/10.1093/scan/nsq005.

Lo, J. H. Y., Fu, G., Lee, K., & Cameron, C. A. (2020). Development of moral reasoning in situational and cultural contexts. *Journal of Moral Education*, *49*(2), 177–193.

McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2022). Moral judgment as categorization (MJAC). *Perspectives on Psychological Science*, *17*(1), 131–152. https://doi.org/10.1177/1745691621990636.

Ochoa, K. D., Rodini, J. F., & Moses, L. J. (2022). False belief understanding and moral judgment in young children. *Developmental Psychology*, *58*(11), 2022–2035. https://doi.org/10.1037/dev0001411.

Parkinson, C., Sinnott-Armstrong, W., Koralus, P. E., Mendelovici, A., McGeer, V., & Wheatley, T. (2011). Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust. *Journal of Cognitive Neuroscience*, *23*(10), 3162–3180. https://doi.org/10.1162/jocn_a_00017.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2014). The role of analytic thinking in moral judgements and values. *Thinking & Reasoning*, *20*(2), 188–214. https://doi.org/10.1080/13546783.2013.865000.

Rhim, J., Lee, G. B., & Lee, J. H. (2020). Human moral reasoning types in autonomous vehicle moral dilemma: A cross-cultural comparison of Korea and Canada. *Computers in Human Behavior*, *102*, 39–56. https://doi.org/10.1016/j.chb.2019.08.010.

Rottman, J., Kelemen, D., & Young, L. (2014). Tainting the soul: Purity concerns predict moral judgments of suicide. *Cognition*, *130*(2), 217–226. https://doi.org/10.1016/j.cognition.2013.11.007.

Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, *11*(2), 233–240. https://doi.org/10.1037/a0022598.

Salerno, J. M., & Peter-Hagene, L. C. (2013). The interactive effect of anger and disgust on moral outrage and

judgments. *Psychological Science*, *24*(10), 2069–2078. https://doi.org/10.1177/0956797613486988.

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288.

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. (2023). afex: Analysis of Factorial Experiments. R package version 1.2-1. https://CRAN.R-project.org/package=afex.

Sunar, D., Cesur, S., Piyale, Z. E., Tepe, B., Biten, A. F., Hill, C. T., & Koç, Y. (2021). People respond with different moral emotions to violations in different relational models: A cross-cultural comparison. *Emotion*, *21*(4), 693–706. https://doi.org/10.1037/emo0000736.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*(10), 780–784. https://doi.org/10.1111/j.1467-9280.2005.01614.x.