

## **RESPONSIBILITY FOR FORGETTING\***

Samuel Murray (Notre Dame), Elise Murray (Tufts), Greg Stewart (Duke)  
Walter Sinnott-Armstrong (Duke), Felipe De Brigard (Duke)

(Forthcoming in *Philosophical Studies*)

### 1. Introduction

Forgetting can be costly. Just ask David Hall, the Springfield, MO fire chief, who responded to a house fire on April 15, 2016 in the 1800 block of North Drury Avenue. The resident was heating some oil for cooking when he stepped outside to check on something. He remembered the oil a short time later—after it ignited and set the kitchen on fire. Hall responded to that house fire, and while the resident suffered minor damage from smoke inhalation, the kitchen was not so lucky. The fire resulted in about \$57,000 of damages, one of the more costly fires that Hall remembered during his tenure as fire chief (Keegan 2016).

Forgetting can also be deadly. Kristie Reeve normally dropped her daughter Sophia off at daycare before working from home. But one summer morning, Kristie had an early meeting and her daughter was sleeping in. This meant Brett Cavaliero—Kristie’s husband—had to take Sophia in before heading to work. Unfortunately, Brett forgot to drop his daughter Sophia off.<sup>1</sup> By the time he remembered, the little girl had been sitting

---

\* We would like to thank Sara Johnson for help with some of the statistical analysis. Also, thanks to Manuel Vargas, Santiago Amaya, Dylan Murray, and Paul Henne for discussion of various points in the paper. Members of MADLab and the Imagination and Modal Cognition lab at Duke University gave critical feedback on the vignettes and analysis of initial results (especially Aaron Ancell, Jesse Summers, Jana Shaich Borg, Luka Ruzic, and Bryce Gessell). Funding for this research was provided through a grant from the Philosophy and Science of Self-Control project through Florida State University (#261898). The views expressed in this paper do not reflect the views of Florida State or the Philosophy and Science of Self-Control project.

<sup>1</sup> Brett, in interviews, never explained why he forgot; but oftentimes when caregivers forget children in cars, there is a familiar pattern of events that occur. The caregiver is normally breaking a habit or routine in taking the child in the car, and usually the caregiver is rushing off to work. See Amaya 2013 for more discussion of these patterns that underlie everyday cases of slips and forgetting.

inside a car, windows closed, on a summer day in Texas for several hours. EMT's declared Sophia dead a mere 80 minutes after Brett called 9-1-1 (Pelletiere 2016).

If you thought that Brett's situation is the sort of monumental tragedy that occurs once every few years, you would be wrong. On average, since 1998, 37 children die *per year* of heat stroke because their parents forget them in hot cars (Weingarten 2009). In 2016, 39 children died of heatstroke in the US (Null 2016). And you also probably underestimated just how often forgetfulness results in house fires. Between 2010-2014, US fire departments responded to 166,100 home fires *per year* that involved some sort of cooking equipment. The National Fire Protection Association's "Home Fires Involving Cooking Equipment" report states that: "Unattended cooking was by far the leading contributing factor in [home] fires and fire deaths," accounting for about 49% of home fires (Ahrens 2016: 39-41). That's a lot of forgetful cooks.

While these stories are dramatic, they might feel strangely familiar. Our lives are full of these sorts of forgetful episodes. And forgetting about something is just one of a multitude of mistakes that we regularly make. For instance, imagine that you are sitting at work when you get a call from your partner asking you to pick something up from the store on your way home. You promise to do so and plan to stop at the store after you leave your office. Now imagine that you get distracted and forget about your plan to go to the store. You drive home without ever thinking about going to the store. When you get home, you realize that you made a mistake. While the details might differ slightly, we have all made mistakes by forgetting to bring about a planned action. And this is the same sort of forgetfulness that plagues the victims of home fires and the parents of infant hyperthermia victims.

In this paper, we focus on forgetting. Specifically, we focus on whether and to what extent we judge that people are responsible for the consequences of their forgetfulness. There is some debate about whether agents are responsible primarily for their *actions* (or omissions) or the states of affairs that obtain in virtue of actions (or omissions) (Frankfurt 1983 and Fischer and Ravizza 1998 defend the former view, while van Inwagen 1978 defends the latter view). In this paper, we presume that agents are responsible primarily for their actions (or omissions) that result from forgetfulness. Thus, in this paper, the phrase ‘responsibility for forgetfulness’ is shorthand for the clunkier, though more precise, phrase ‘responsibility for actions (or omissions) that result from forgetfulness’.

We focus on cases of forgetfulness for three reasons. First, these cases of forgetfulness constitute one interesting class of unwitting omissions, and philosophers have recently been interested in the conditions under which individuals might be responsible for some of their unwitting omissions (Clarke 2016; Guerrero 2007; Harman 2011; Shoemaker 2011; Levy 2009; Smith 2005; Smith 1983; Amaya and Doris 2014).

Second, there is a substantial debate over what factors might mitigate responsibility for forgetfulness. For instance, does it matter to people whether someone cares a lot about doing the thing that they forget to do? Does it matter whether someone forgets about something when they are under a lot of stress? Getting clear on these issues might help to determine what to do with people whose forgetfulness results in fire, death, or some other negative outcome—a question that vexes legal scholars working on issues of culpability and liability for negligence (Husak 2011; Hurd and Moore 2011; Yaffe 2012; Collins 2006; Finkel and Groscup 1997: 70).

Finally, explaining responsibility for these unwitting omissions cases has important implications for a long-running debate about the nature of responsible agency (more on this shortly).

Getting clear on these questions requires settling some background issues, such as understanding how people actually assign responsibility for forgetfulness. Unfortunately, to the best of our knowledge, there has been no systematic study of our judgments about whether and in what circumstances people hold others responsible for forgetfulness. The studies reported in the current paper begin to fill this lacuna.

In particular, this information is relevant to a debate about the nature of responsible agency. Broadly speaking, there are two different theories of what makes individuals apt targets for various culpability-imputing attitudes: capacitarian and valuational theories. Capacitarian theories state that agents are responsible for some outcome *O* in virtue of certain responsibility-relevant capacities figuring (causally) in the production of some behavior that results in *O*'s obtaining (cf. Wolf 1990, Fischer and Ravizza 1998, Nelkin 2008, Vargas 2013). According to valuational theories what makes some individual responsible for some outcome, *O*, is that the behavior that results in the *O*'s obtaining has a particular valuative psychological state as part of its etiology (Frankfurt 1971, Watson 1975, Smith 2005, Doris 2015, Sripada 2016). Some valuational theories state that the relevant valuative state is caring (Sripada 2015, 2016; Björnsson 2017; Smith 2005; Mason 2015). In this paper, we are interested in the caring kind of valuative state.<sup>2</sup>

---

<sup>2</sup> It is an open question whether care is just one kind of valuative state relevant to responsible agency or whether care is equivalent to the whole class of valuative states relative to responsible agency. Some philosophers disagree about the nature of caring itself. For instance, is caring just the relation of desire (as in Arpaly and Schroeder 2014)? Or is caring a more complex, interconnected group of dispositions to be in certain emotional states and exhibit certain patterns of attentional focus (as in Jaworska 1999)? There are

Cases of forgetfulness present an interesting split between valuational and capacitarian theories because the two theories make different predictions about responsibility in these cases. In cases of forgetfulness (like the ones we present below), agents seem to be responsible for their forgetfulness and two things seem true:

- (1) Forgetful agents do not fulfill standard valuationist conditions on responsibility,  
and;
- (2) Forgetful agents do seem to fulfill standard capacitarian conditions on  
responsibility.

Some evidence for (1) comes from the fact that valuationists treat agents in cases of forgetfulness (or structurally similar cases) as not responsible for the consequences of their forgetfulness (Sripada 2015: 261).<sup>3</sup> Other valuational theorists, like John Doris, concede that the valuational theory cannot easily explain responsibility for forgetfulness. But this is not too troubling, on Doris' view, because these cases of forgetfulness are rare (Doris 2015: 154-55). Valuational theorists, then, either admit that agents are not responsible for the consequences of their forgetfulness *or* try to dismiss intuitions that forgetful agents are sometimes responsible for their forgetfulness as irrelevant for theory construction.

Capacitarian theorists, on the other hand, agree that forgetful agents fulfill some suitably

---

other issues (raised in Darwall 2002; Frankfurt 1999; Sripada 2016), but we set those aside because substantive disputes about the nature of caring are orthogonal to the present discussion.

<sup>3</sup> Sripada (2015: 261) provides a case called *Mrs. Smith—Forgets*: “Mrs. Smith is playing cards with her friends and someone happens to mention today’s date. Unfortunately, this doesn’t trigger Mrs. Smith’s remembering that it is her grandson’s birthday. As a result, she does not give him a call even though she has done this in year’s past. Billy is sad.” Sripada continues with his own diagnosis of the case, claiming: “*Mrs. Smith—Forgets* is, I believe, a case where the ability to express one’s self over a suitably wide range of cases is preserved, but the manifestation of the ability in the actual circumstances is compromised; that is, her self fails to be expressed in what she does in the actual situation. Moreover, I believe that intuitively she is not morally responsible for her forgetting Billy’s birthday.” In the terminology of this paper, Sripada concedes that Mrs. Smith fulfills the capacitarian conditions on responsibility (having the ability to express herself), but she fails to express herself in forgetting (thereby not fulfilling a certain version of the valuationist theory of responsible agency).

specified conditions on responsibility (even though the details of various theories might diverge; see Clarke 2014, Murray 2017, McGeer and Pettit 2015, Timpe 2011).

Thus, these cases present instances where the two main theories of the nature of responsible agency offer different predictions about cases. So, the patterns of judgment that people exhibit in response to these cases would provide strong evidence in favor of one theory over another. These patterns of judgment are important because both theories attempt to explain folk thinking about responsible agency. If either theory had to adopt a revisionist or eliminativist position, it would be a serious cost for that theory. So the judgments that people exhibit in response to these cases constitute an important class of evidence for theories of responsible agency. If people do not judge that agents are responsible for their forgetfulness, then this would be evidence for a valuational theory. But if people judge that agents are responsible for their forgetfulness, it suggests that our folk practices of assigning responsibility recognize that agents can be responsible even when there is not some relevant valuative state (or set of valuative states) that is causally related to the behavior that brings about the outcomes for which the agent is responsible. This latter scenario would be evidence for a capacitarian theory and evidence against a valuational theory. Additionally, if people are disposed to judge that agents are responsible for forgetfulness, this would be problematic for the valuational theory. Many valuationists tend to marshal support for their theory by pointing to the pretheoretical credibility of the theory (Woolfolk et al. 2006; Sripada 2015: 263; Smith 2005; Björnsson 2017). If the valuational theory conflicts with some established feature of our folk practices of assigning responsibility, this would undercut a major motivation for adopting the valuational view.

In the current studies, we focus exclusively on judgments of responsibility in cases where an agent forgets to perform some planned action. Other studies have focused on judgments of responsibility for phenomena that are similar to forgetfulness, like acting in light of a false belief or intending to do something good but accidentally bringing about a bad outcome (Martin and Cushman 2016; Young, Nichols, and Saxe 2010; Young and Tsoi 2013). Forgetfulness, however, is neither an accident nor a doxastic mistake; rather, forgetfulness seems to lie somewhere between these two.

A mistake usually involves acting with a false belief. For example, if I step on your foot, but I believe (falsely) that your foot is just another part of the subway car, then my stepping on your foot is a mistake (think: “Oh, that was your foot? My mistake”).<sup>4</sup> Being forgetful, however, does not require having false beliefs. At most, when one is forgetful, one has dispositions to believe true things, but these dispositions fail to manifest. Thus, being forgetful implies a lack of true beliefs, but not necessarily having false beliefs. Forgetfulness is not simply an accident, either. Accidents are trickier to characterize, though accidents typically involve some sort of involuntary or unintentional behavior, and that the source of this involuntary or unintentional behavior cannot be traced back to an exercise (or instance of) intentional agency. For instance, if I step on your foot because the subway lurched forward abruptly, my stepping on your foot is an accident only if there is no exercise (or instance of) intentional agency on my part that explains or is causally implicated in my stepping on your foot. Forgetfulness does not necessarily

---

<sup>4</sup> There are two dimensions along which this issue gets complicated. The first has to do with the relationship between false belief, credence, and fragmented belief (e.g., do I have a false belief with respect to your foot if I have a credence of .12 that your foot occupies the part of the subway car on which I’m about to step?). Second, do mistakes ever involve culpable false belief? Suppose that there is such a thing as culpable belief (another complicated issue). If I believe (falsely) that your foot is part of the subway car, and I am culpable for believing falsely, then do I still make a mistake in stepping on your foot? Thankfully, we can sidestep these complications in our discussion.

involve involuntariness or unintentional behavior. One could still be acting intentionally (or there could be at least one description of one's behavior under which one is acting intentionally) even when one is forgetful (see Amaya 2013: 569). Thus, our study of responsibility for forgetfulness is distinct from studies that focus on mistaken or accidental wrongdoing.

We examine four hypotheses in the current study:

- (H1) People will judge that agents are sometimes responsible for their forgetfulness.
- (H2) These judgments will be sensitive to the level of stress that the agent experiences while being forgetful.
- (H3) Stress will provide a partial excuse, but it will not function as a total excuse for some of the consequences of being forgetful.
- (H4) Judgments of responsibility will not be sensitive to how much the agent cares about performing the forgotten action.

We posit hypotheses (H2)-(H4) because we believe that these align with the prediction of a capacitarian theory of responsibility. These hypotheses also seem to be those to which the folk are pretheoretically committed. This latter point is based on an argument recently put forward by Manuel Vargas (2013: 137-52) the conclusion of which states that the capacitarian theory best explains current psychological research on folk responsibility judgments *and* that apparent counterevidence can be explained away. We will not rehearse his argument here, but we take it as evidence for our hypotheses (H2)-(H4).

Our choice of stress and care as dependent variables derives from the debate between valuational and capacitarian theories of responsibility. Care picks out the range of psychological states that many valuationists take to be relevant to responsible agency. As stated above, valuationists take hold that an agent is responsible for her behavior when that behavior reflects some privileged subset of valuative states of the agent (differences between theories reflect differences about what counts as the privileged subset). Care

includes those pro-attitudes that are constitutive of the agent's true or *deep* self (cf. Sripada 2016: 1208; Smith 2005: 251; De Freitas et al., 2017). When the agent behaves in a way that is an apt target for moral evaluation, the agent's behavior expresses or manifests (at least some subset of) these cares. When the agent behaves in a way that fails to express or manifest some care (either because the behavior is coerced, pathological, etc.), then the agent (or the agent's behavior) is not an adequate candidate for moral evaluation. Thus, the valuational theory predicts that the absence of care should result in judgments of diminished responsibility.<sup>5</sup>

Stress is relevant for the capacitarian theory because stress affects capacity. Stress affects executive function (Scult 2017; Sapolsky 2015; McEwen 2015 et al.; Hollon et al. 2015; Arnsten 2015; Goldfarb et al. 2017), thereby making it more difficult (or practically impossible) to exercise one's responsibility-relevant capacities. In this paper, we assume that an agent is under stress only if an agent's attention is drawn away from some immediate or ongoing task to some present or future event, where:

- (i) The occurrence of the event prompts some response from the agent (or, in the case of future-directed stress, the consideration of the event is connected to dispositions to reflect on what sorts of responses *will* be appropriate when the event occurs).
- (ii) The occurrence of the event results in the agent's having some fitting affective response (or, in the case of future-directed stress, the

---

<sup>5</sup> This picture is complicated by the fact that absence of care does not always mitigate. This is because sometimes, an agent may occupy a role or be under an obligation that requires her to have certain cares. In these cases, the valuationist would not predict that absence of care results in reduced responsibility (e.g., a parent is not excused for neglecting her kids simply because she doesn't care about them). We have chosen vignettes that do not run into this complication. That is, in the vignettes we present below, it is not the case that the agent is required to care to some degree about the action that he forgets to perform (it *is* the case that the agent is required to *do* the action, but this does not entail that the agent must *care* about the action).

consideration of the event is connected to dispositions to come to have some fitting affective response to the prospect of experiencing the event).<sup>6</sup>

- (iii) The occurrence of the event deviates from some expectation the agent has about what should or should not be happening (or, in the case of future-directed stress, certain features of the event are unknown and unpredictable).<sup>7</sup>

The degree of stress that some agent is under is a function of how much effort is needed to draw attention away from some present or future event back to an immediate or ongoing task. On this understanding of stress, stress is not essentially tied to a particular kind of affective response to the present or future event (though there must be some fitting affective response to some present or future event in order for there to be stress), nor does stress require that the present or future event to which attention is drawn has a certain negative valence (notice also that we propose merely necessary conditions for an agent being under stress). Thus, the capacitarian theory predicts that the presence of (a certain degree of) stress should result in judgments of diminished responsibility.<sup>8</sup>

In our studies, rather than just measuring judgments of responsibility directly, we measured judgments of fault, appropriate blame, and guilt. We did this for two reasons. First, the term ‘responsibility’ is ambiguous. One kind of responsibility is causal responsibility, while another is liability responsibility (Hart 1968; Raz 2011: 227-229).

---

<sup>6</sup> These fitting affective responses will sometimes have a distinctively negative valence (as in cases of negative stress), though we leave open the possibility of positive stress, so that fitting affective responses require the agent’s coming to be in some positively valenced affective state. See Maxwell and Racine 2012 for more on the idea that stress requires some affective component.

<sup>7</sup> For example, if someone is stressed about an upcoming performance, the cause of the stress might be, in part, the fact that one is unsure about how the performance will go *and* the fact that one is incapable of resolving this uncertainty until the event occurs. The role of uncertainty in stress matches Reynolds, Owens, and Rubenstein’s (2012) definition of stress as a psychological state grounded in uncertainty about fulfilling certain duties or roles.

<sup>8</sup> Further issues about the scope of stress are discussed in section 4.2.

Additionally, philosophers often distinguish among three different senses of liability responsibility, namely accountability, answerability, and attributability (Watson 1996; Smith 2008; Hieronymi 2008; Shoemaker 2015). We are interested in the accountability sense of liability responsibility, but because ‘responsibility’ is ambiguous in this multi-faceted way, measuring judgments of responsibility directly could be confusing (Malle, Guglielmo, and Monroe 2012).

This might suggest focusing our measurements on distinctively *moral* responsibility, thus isolating the accountability sense of liability responsibility that interests philosophers and psychologists. This leads to our second reason for focusing on fault, blame, and guilt, namely that ‘moral responsibility’ is a term of art. It is unclear that people use the term ‘moral responsibility’ in ordinary conversational contexts, and so unclear whether there is a folk concept of moral responsibility to probe or even whether people have a coherent concept of moral responsibility (see van Inwagen 2017: 209). So while measuring judgments of moral responsibility would avoid the ambiguity of judgments of responsibility, it is not obvious whether there is anything to measure.

‘Fault’, ‘guilt’, and ‘blame’ are all terms of ordinary English, and so we are more confident that English speakers possess the normative concepts of fault, guilt, and blame. That is, the folk might have concepts of fault, guilt, and blame that guide and constrain responsibility ascriptions. At this point, someone might wonder why we invoke the concept of responsibility at all. Why not focus exclusively on the normative concepts of fault, guilt, and blame? The reason for this is mostly conceptual. Different philosophers have proposed that ‘S is morally responsible for  $x$ ’ (where  $x$  names an event or a state of affairs) just means either: (a) S is at fault for  $x$ ’ing or bringing it about that  $x$  (van Inwagen 2008: 328); (b) It is appropriate for someone to blame S for  $x$ ’ing or bringing it about that

$x$  (Rosen 2015: 66-67), or; (c) S should feel guilty about  $x$ 'ing or bringing it about that  $x$  (Strawson 2003: 85). Thus, on this conceptual scheme, probing judgments of fault, blame, and guilt just is a way of probing judgments about moral responsibility, even if the folk lack a theory of moral responsibility that makes explicit the interrelations between these different concepts. For these reasons, we focus on fault, appropriate blame, and guilt.

## 2. Study 1

### 2.1 Method

285 participants were recruited through Amazon's Mechanical Turk (Age  $M=33.98$ ,  $SD=10.53$ ). Participants were made aware that their answers would be anonymous, and were monetarily compensated for their participation (\$.50 for an approximately two minute task). Four participants failed an attention check, and were thus removed from the data set prior to analysis (final  $N=281$ ). Participants were asked to read a vignette (see below) describing a situation where an agent forgets to bring about some planned action. The situation is similar to examples of morally significant forgetfulness used in the philosophical literature (cf. Clarke 2014; Mele 2006; Murray 2017).

We manipulated two variables in the vignettes: (a) the level of care that the agent displays toward performing the planned action and (b) the level of stress that the agent is under when the agent needs to (and has the opportunity to) initiate performance of the planned action. Thus, we designed four vignettes that correspond to the four possible combinations of care and stress: High Care, High Stress (HCHS;  $n=62$ ); Low Care, High Stress (LCHS;  $n=74$ ); High Care, Low Stress (HCLS;  $n=70$ ), and; Low Care, Low Stress (LCLS;  $n=78$ ). Participants were randomly assigned to one of these four vignettes. The vignette we used was as follows (for clarity, we list the HCHS and LCLS vignettes

separately. The ‘care’ element is underlined and the ‘stress’ element of the vignette is italicized)<sup>9</sup>:

(HCHS) Randy’s wife calls him and asks him to pick up some ingredients from the store on the way home from work. Randy is hosting a birthday party for his close friend today and his wife needs the ingredients to bake the birthday cake. The cake requires special ingredients that are only available at one grocery store in the area. The store is on the route that Randy takes to get home and Randy wants to get the ingredients for the cake, so he tells his wife that he’ll pick up the ingredients on the way home. Just before he gets in his car, *Randy gets a call from someone that works at his office telling him that he very likely will not get the promotion he was hoping for. The news troubles Randy, because he needed the additional money to help with his mortgage payment. Randy gets nervous and starts to think about the different things he could do at work and other ways to bring in extra money to help with the payments.* Because of this, Randy doesn’t notice that he is approaching the store. He drives right by the store and arrives home without the ingredients. Because the store is a specialty store, it closes early in the evening. By the time Randy gets home, the store has closed and it is too late for him to turn around.

(LCLS) Randy’s wife calls him and asks him to pick up some ingredients from the store on the way home from work. Randy is hosting a birthday party for his close friend today and his wife needs the ingredients to bake the birthday cake. The cake requires special ingredients that are available at only one store in the area. The store is on the route that Randy takes to get home. However, Randy doesn’t like that particular store. It’s always crowded and the cashiers are mean to him. He doesn’t want to go, but he tells his wife that he’ll pick up the ingredients on the way home. Just before he gets in his car, *Randy gets a call from someone that works at his office telling him that he very likely will get the promotion he was hoping for. The news pleases Randy, because he needs the additional money to buy some things he wants. Randy gets excited and, because of this, he doesn’t notice that he is approaching the store. He drives right by the store and arrives home without the ingredients. Because the store is a specialty store, it closes early in the evening. By the time Randy gets home, the store has closed and it is too late for him to turn around.*

Participants were asked to indicate whether they agreed or disagreed with the following statements: (1) “It is Randy’s fault that he did not get the ingredients.” (2) “It is fair for Randy’s wife to blame him for not getting the ingredients.” (3) “Randy should feel guilty

---

<sup>9</sup> In the HCLS and LCHS vignettes, we substituted the appropriate care or stress element into the same spot in the vignette to maintain relative symmetry between the vignettes.

for not getting the ingredients.” Participants rated their answers on a 7-point Likert scale, ranging from 1 = *highly disagree* to 7=*highly agree*. Additionally, participants were asked a question about frequency: “How common do you think this kind of mistake is?” using a 7-point Likert scale, ranging from 1 = *very rare* to 7 = *very common*. The first three statements were presented in random order, and these were always followed by the frequency question.

## 2.2 Results

In order to assess statistical significance of care and stress on the outcome variables of fault, blame, and guilt scores, responses were modeled in a 2 (high/low care) by 2 (high/low stress) between-subjects analysis of variance (ANOVA) for each dependent variable. Analyses were conducted using IBM SPSS 22.0 and statistical significance was measured at  $\alpha = .05$ . See Table 1 for all means and standard deviations.

For fault, there was a significant main effect of both care ( $F(1,280) = 7.768, p < .01$ , partial  $\eta^2 = .027$ ) and stress ( $F(1,280) = 5.350, p < .05$ , partial  $\eta^2 = .019$ ), and no significant interaction. This indicates that fault ratings are higher in LC and LS conditions relative to HC and HS conditions.

For blame, there was a significant main effect of both care ( $F(1,280) = 9.521, p < .01$ , partial  $\eta^2 = .033$ ) and stress ( $F(1,280) = 14.125, p < .001$ , partial  $\eta^2 = .048$ ), and no interaction. This indicates that blame ratings are higher in LC and LS conditions relative to HC and HS conditions.

Finally, for guilt, there was a main effect of care ( $F(1,280) = 4.628, p < .05$ , partial  $\eta^2 = .016$ ), and a marginal main effect of stress ( $F(1, 280) = 3.719, p = .055$ , partial  $\eta^2$

=.013), and no interaction. This indicates that guilt ratings are higher in LC and LS conditions relative to HC and HS conditions.

Finally, participants gave similar answers to one another across conditions about the commonality of the mistake depicted in the vignette (HCHS  $M=5.26$ ,  $SD=1.4$ ; LCHS  $M=5.68$ ,  $SD=1.2$ ; HCLS  $M=5.46$ ,  $SD=1.4$ ; LCLS  $M=5.32$ ,  $SD=1.4$ ). The mean ratings of these perceptions did not differ significantly ( $F(3,280) = 1.342$ ,  $p = .26$ )

### 2.3 Discussion

These results suggest that participants were more likely to agree that it was Randy's fault, that it was fair for his wife to blame him, and that he should feel guilty for not getting the ingredients when Randy both does not care much about performing the action *and* also when his stress levels are perceived to be low. Conversely, participants were less likely to say that it was his fault, that his wife was fair in blaming him, and that he should feel guilty for not getting the ingredients when Randy both cares about performing the action and was under conditions of high stress.

The effect sizes of these variables, however, are small-to-medium ( $\eta^2 < .14$ ). We think that four different factors might explain these effect sizes. First, in the high stress conditions, Randy's stress results from his not having the money to pay his mortgage in virtue of his getting passed over for a promotion. Participants, however, either might not interpret this event as stressful *or* they might evaluate Randy in a negative light because he took out a mortgage expecting to get money through a promotion (an imprudent move that might lead participants to blame Randy in light of faulty or sub-optimal character).

Second, in the high care conditions no explanation was given as to why Randy cares about getting the ingredients. This is problematic because the subjects might

wonder how much Randy really cares. This creates an additional (third) problem, namely that participants might misconstrue the level of care that Randy exhibits in the high care conditions.

Finally, the questions that we ask are ambiguous in a crucial way. We asked participants about the degree to which they agreed with statements about whether Randy is at fault, would be appropriately blamed, and should feel guilty. But this can be interpreted in one of two ways. One interpretation treats these properties dichotomously, so that Randy either should or should not feel guilty (for example). Another interpretation treats these properties as scalar, so that Randy should feel guilty *to some degree*. Our questions do not disambiguate between these, so that a high rating might signal agreement with the dichotomous interpretation without giving any indication about the scalarity of the property.

Additionally, we did not include a manipulation check to determine whether participants were accurately interpreting Randy's level of care across different conditions. For instance, someone might suggest that participants in this study were judging that Randy is blameworthy and *then* judging that Randy's level of care is low. There is no way to ensure that blameworthiness judgments are not driving assessments of care. To correct these limitations, and to corroborate the effects from Study 1, we conducted a second study.

### 3. Study 2

#### 3.1 Method

420 participants were recruited through Amazon's Mechanical Turk (Age  $M=33.69$ ,  $SD=10.96$ ). We removed 21 participants from analysis who failed a more rigorous

attention check (final  $N=399$ ). The methodology of Study 2 is identical to that of Study 1, except that vignettes in Study 2 conveyed better information about the level of care. In the High Care condition, for instance, the vignettes in Study 2 explain why Randy cares about getting the ingredients, rather than simply stating that he cares about getting the ingredients. In the Low Care condition, we changed the information about the source of the low care to something that pertains to the cake and getting the ingredients.

Additionally, the vignettes in Study 2 depicted stressors that participants could not (reasonably) interpret as the agent's fault. We used the same combination of care and stress to generate four vignettes: High Care, High Stress (HCHS;  $n=94$ ); Low Care, High Stress (LCHS;  $n=91$ ); High Care, Low Stress (HCLS;  $n=104$ ), and; Low Care, Low Stress (LCLS;  $n=110$ ). The specific vignette used was the following (again, we list the HCHS and LCLS vignettes separately):

(HCHS) Randy's wife calls him and asks him to pick up some ingredients from the store on the way home from work. Randy is hosting a birthday party for his close friend today and his wife needs the ingredients to bake the birthday cake. The cake requires special ingredients that are available at only one store in the area. The store is on the route that Randy takes to get home. Randy really cares about getting the ingredients. His friend doesn't care about birthday cakes, but Randy thinks that baking a birthday cake is very important. Because he cares about it, Randy wants to get the ingredients for the cake, so he tells his wife that he'll pick up the ingredients on the way home. Just before he gets in his car, *Randy gets a call from his supervisor telling him that he will get laid off at the end of the week due to downsizing. The news troubles Randy, because without his job he won't be able to pay his rent and will lose his health insurance. Randy gets nervous and starts to think about different jobs that he could apply to.* Because of this, Randy doesn't notice that he is approaching the store. He drives right by the store and arrives home without any of the ingredients. Because the store is a specialty store, it closes early in the evening. By the time Randy gets home, the store has closed and it is too late for him to turn around.

(LCLS) Randy's wife calls him and asks him to pick up some ingredients from the store on the way home from work. Randy is hosting a birthday party for his close friend today and his wife needs the ingredients to bake the birthday cake. The cake requires special ingredients that are available

at only one store in the area. The store is on the route that Randy takes to get home. However, Randy doesn't really care about getting the ingredients. He and his wife have already made a big effort to plan a great party, so having a cake doesn't seem very important. Despite the fact that he doesn't care about going, he tells his wife that he'll pick up the ingredients on the way home. Just before he gets in his car, *Randy gets a call from someone that works at his office telling him that he very likely will not get the promotion he was hoping for. Randy gets upset at this, because without the promotion he won't get the extra money that he needs to fund a vacation that he was planning. Randy starts to think about additional things that he could do at work to bring in extra money.* Because of this, Randy doesn't notice that he is approaching the store. He drives right by the store and arrives home without any of the ingredients. Because the store is a specialty store, it closes early in the evening. By the time Randy gets home, the store has closed and it is too late for him to turn around.

In addition to altering the vignette, we altered the questions that participants answered. Instead of asking for the degree of agreement or disagreement with particular statements, we asked participants the following questions: (1) How much does Randy care about getting the ingredients for the cake? (2) How much fault does Randy have for not getting the ingredients? (3) How much should Randy's wife blame him for not getting the ingredients? (4) How much guilt should Randy feel about not getting the ingredients? (5) How much moral responsibility does Randy have for not getting the ingredients? (6) How common is the mistake that Randy made? All of these questions were presented in random order to reduce the chances of any ordering effects. We added the question about care to ensure that participants were accurately interpreting the level of care that Randy displays. Also, we added the question about moral responsibility to see whether participants' assessments of moral responsibility correlated with their assessments of fault, blame, and/or guilt. This would enable us to test various analyses of responsibility (mentioned in the introduction) that posit a conceptual connection between moral responsibility and fault, blame, and guilt.

### 3.2 Results

Similarly to Study 1, we assessed statistical significance of care and stress on the dependent variables of fault, blame, guilt, and moral responsibility scores, by modeling a 2 (high/low care) by 2 (high/low stress) between-subjects analysis of variance (ANOVA) for each dependent variable. See Table 2 for all means and standard deviations.

For fault, we found significant main effects of care ( $F(1,395) = 5.446, p < .01$ , partial  $\eta^2 = .014$ ) and stress ( $F(1,395) = 38.183, p < .001$ , partial  $\eta^2 = .088$ ) and no interaction. This indicates that fault ratings are higher in LC and LS conditions relative to HC and HS conditions.

For blame, there were significant main effects of care ( $F(1,395) = 14.038, p < .001$ , partial  $\eta^2 = .034$ ) and stress ( $F(1,395) = 70.746, p < .001$ , partial  $\eta^2 = .152$ ), and no interaction. This indicates that blame ratings are higher in LC and LS conditions relative to HC and HS conditions.

For guilt, we found significant main effects of care ( $F(1,395) = 6.914, p < .01$ , partial  $\eta^2 = .017$ ) and stress ( $F(1,395) = 50.409, p < .001$ , partial  $\eta^2 = .113$ ), and no interaction. This indicates that guilt ratings are higher in LC and LS conditions relative to HC and HS conditions.

For moral responsibility, there were significant main effects for both care ( $F(1,395) = 8.172, p < .01$ , partial  $\eta^2 = .020$ ) and stress ( $F(1,395) = 20.048, p < .001$ , partial  $\eta^2 = .048$ ), and no interaction. This indicates that moral responsibility ratings are higher in LC and LS conditions relative to HC and HS conditions.

Participant assessments of Randy's level of care gravitated toward the extremes. The modal response in the low care conditions was 1 (60% of participants in LCHS gave a care rating of 1, while 52.7% of participants in LCLS gave a care rating of 1), whereas

the modal response in the high care conditions was 7 (45.7% of participants in HCHS gave a care rating of 7, while 51.9% of participants in HCLS gave a care rating of 7). These care assessment distributions are in the expected directions based on the corresponding care condition.

Finally, to verify whether participants correctly interpreted the level of care that Randy displays in the vignette, we analyzed the ratings of care. Participants' rating of care in the low care conditions (LCLS  $M=1.71$ ,  $SD=1.2$ ; LCHS  $M=2.33$ ,  $SD=1.8$ ) was significantly lower than in the high care conditions (HCHS  $M=5.86$ ,  $SD=1.5$ ; HCLS  $M=5.80$ ,  $SD=1.6$ ;  $F(3, 398) = 210.451$ ,  $p < .001$ ). Direct comparisons between mean ratings of care showed that ratings of care in the LCHS condition were significantly higher than ratings of care in the LCLS condition ( $p < .05$ ), and ratings of care in the HCLS and HCHS conditions were significantly higher than ratings of care in the LCHS condition ( $p < .001$  in both cases). Additionally, participants gave similar answers to one another across different conditions as to how common the kind of mistake depicted in the vignette is (HCHS  $M=4.65$ ,  $SD=1.4$ ; LCHS  $M=5.01$ ,  $SD=1.3$ ; HCLS  $M=4.95$ ,  $SD=1.5$ ; LCLS  $M=4.71$ ,  $SD=1.7$ ). The mean ratings of these assessments did not differ significantly ( $F(3, 398) = 1.453$ ,  $p = .23$ ).

### 3.3 Post-hoc Comparison of Care and Stress Ratings by Study

We found significant main effects of both care and stress across studies and models (except for stress on guilt in Study 1), which indicates that care and stress were both significant predictors of blame, fault, and guilt ratings. Although both factors were significant, each independent variable had varying effect sizes in each model specified. The effect sizes of care and stress were small in Study 1 and bigger (though still relatively

small) in Study 2. Noticeably, however, the effect sizes of care on judgments of blame and guilt were essentially the same (Blame: partial  $\eta^2 = .033$  in Study 1, partial  $\eta^2 = .034$  in Study 2; Guilt: partial  $\eta^2 = .016$  in Study 1 and partial  $\eta^2 = .017$  in Study 2), whereas the effect size of care on judgments of fault were fairly close, but not the same, partial  $\eta^2 = .027$  to partial  $\eta^2 = .014$  between studies). On the other hand, the effect size of stress differed greatly between studies. As a result of these effect size differences, we hypothesized, post-hoc, that there may be a significant difference in how participants are rating stress, as well as a difference in effect size, based on the study in which they participated (i.e., 1 or 2), and that there would not be a significant difference in participant ratings and effect sizes of care based on study.

In order to test this post-hoc hypothesis, we modeled the three ANOVA models with Fault, Blame, and Guilt as dependent variables, with Study as an additional factor. The Study variable was also specified to interact with the two original independent variables, Care and Stress. Significant interaction terms of Care\*Study and/or Stress\*Study would indicate that Study participation significantly affected care and/or stress ratings; as such, those interaction statistics are reported below.

For fault, blame, and guilt, the Care\*Study interaction variable was not statistically significant and the effect size of the interaction term (that is, the effect size change from Study 1 to Study 2) was negligible (partial  $\eta^2 < .001$ ), whereas the Stress\*Study interaction variable was statistically significant with small effect sizes (Fault:  $F(1, 674) = 8.110, p < .01$  partial  $\eta^2 = .012$ ; Blame:  $F(1, 674) = 7.677, p < .01$  partial  $\eta^2 = .011$ ; Guilt:  $F(1, 674) = 7.457, p < .01$  partial  $\eta^2 = .011$ ). These results indicate that there was no noticeable change in effect size of care or stress based on study.

### 3.4 Discussion

As in Study 1, the results of Study 2 suggest that participants were more likely to judge that Randy is more at fault, that his wife can fairly blame him more, that he should feel more guilt, and that he is more morally responsible for not bringing the ingredients when he does not care much about getting the ingredients, and also when his stress levels are low. Conversely, participants were more likely to judge that Randy is less at fault, that his wife can fairly blame him to a limited degree, that he should feel less guilt, and that he is less morally responsible for not getting the ingredients when Randy cares about performing the action and is also under high stress.

## 4. Study 3

While Study 2 replicates the findings in Study 1, both Studies rely on the same vignette. Additionally, both vignettes are such that the care element pertains to family while the stress element pertains to work. In order to show that our results are not an artifact of the vignette, we decided to run another study with different vignettes. In these vignettes, the care element pertains to work while the stress element pertains to family.

### 4.1 Method

400 participants were recruited through Amazon's Mechanical Turk (Age  $M=35.3$ ,  $SD=10.7$ ). We removed 13 participants from analysis who failed a rigorous attention check (final  $N=387$ ). The methodology of Study 3 is identical to that of Studies 1 & 2, except that we use different vignettes in Study 3 that have different care and stress elements. We used the same combination of care and stress to generate four vignettes: High Care, High

Stress (HCHS;  $n=97$ ); Low Care, High Stress (LCHS;  $n=99$ ); High Care, Low Stress (HCLS;  $n=93$ ), and; Low Care, Low Stress (LCLS;  $n=98$ ). The specific vignette used was the following (again, we list the HCHS and LCLS vignettes separately):

(HCHS) Whenever Randy needs to travel for work, he needs to submit a Travel Request Form online to get the expenses approved. Without the Form, Randy's company cannot authorize reimbursements for work-related travel. Randy needs to submit a form by the end of the day to get approval for traveling to a work conference next week. Randy's boss wants him to go to this conference. Randy cares a lot about going to the conference. He's hoping to meet with a number of potential investors who might invest good money in Randy's company. Because of this, and because his boss wants him to, Randy plans to go to the conference. As Randy is preparing the form, *he gets a call from his brother that their father had to be admitted to the hospital for severe chest pain. The news troubles Randy because his dad is old and is in bad health. Randy gets nervous and starts looking on the internet for information about what his father's condition might be.* Because of this, Randy forgets about the form. By the time he remembers, it's too late to submit the form. This means that Randy will not be able to travel to the conference, as he cannot afford to pay the travel expenses out of pocket.

(LCLS) Whenever Randy needs to travel for work, he needs to submit a Travel Request Form online to get the expenses approved. Without the Form, Randy's company cannot authorize reimbursements for work-related travel. Randy needs to submit a form by the end of the day to get approval for traveling to a work conference next week. Randy's boss wants him to go to this conference. Randy doesn't really care about going to the conference. Randy doesn't really need to network with anybody going to the conference, and the sessions don't seem interesting. However, because his boss wants him to, Randy plans to go to the conference. As Randy is preparing the form, *he gets a call from his brother. Randy he gets a call from his brother. Randy had gotten a pair of football tickets for Saturday and invited his brother. Unfortunately, something came up and his brother had to cancel. Randy gets distracted thinking about whom he could give the extra ticket to on short notice or whether he should just try to sell the tickets.* Because of this, Randy forgets about the form. By the time he remembers, it's too late to submit the form. This means that Randy will not be able to travel to the conference, as he cannot afford to pay the travel expenses out of pocket.

We asked participants the same questions as in Study 2, substituting the appropriate omission (i.e., Randy's not submitting the form) and changing the question about blame to: How much should Randy's boss blame him for not submitting the form?

## 4.2 Results

Similarly to Studies 1 & 2, we assessed statistical significance of care and stress on the dependent variables of fault, blame, guilt, and moral responsibility scores, by modeling the data as a 2 (high/low care) by 2 (high/low stress) between-subjects analysis of variance (ANOVA) for each dependent variable. See Table 3 for all means and standard deviations.

For fault, there was no significant main effect of care ( $F(1,383) = 3.465, p = .063$ , partial  $\eta^2 = .009$ ), though there was a significant main effect of stress ( $F(1,383) = 157.968, p < .001$ , partial  $\eta^2 = .292$ ) and no interaction. This indicates that fault ratings are higher in the low stress conditions and lower in the high stress conditions.

For blame, again, there was no significant main effect of care ( $F(1,383) = 1.175, p = .279$ , partial  $\eta^2 = .003$ ), though there was a significant main effect of stress ( $F(1,383) = 215.762, p < .001$ , partial  $\eta^2 = .360$ ), with no interaction. This indicates that blame ratings are higher in the low stress conditions and lower in the high stress conditions.

For guilt, there were significant main effects of care ( $F(1,383) = 11.534, p < .001$ , partial  $\eta^2 = .029$ ) and stress ( $F(1,383) = 160.677, p < .001$ , partial  $\eta^2 = .296$ ), and no interaction. This indicates that guilt ratings are higher in LC and LS conditions relative to HC and HS conditions.

For moral responsibility, we found significant main effects for both care ( $F(1,383) = 4.731, p < .03$ , partial  $\eta^2 = .012$ ) and stress ( $F(1,383) = 70.297, p < .001$ , partial  $\eta^2 = .155$ ), and no interaction. This indicates that moral responsibility ratings are higher in LC and LS conditions relative to HC and HS conditions.

Care assessment distributions were in the expected directions based on the corresponding care conditions. In the low care conditions, most participants gave a care rating of 1 or 2 (62.3% of participants in the LCLS condition and 65.6% of participants in the LCHS condition). Most participants in the high care conditions gave a care rating of 6 or 7 (53.8% in the HCLS condition and 60.8% in the HCHS condition).

To verify whether participants correctly interpreted the level of care that Randy displays in the vignette, we analyzed the ratings of care. Participants' ratings of care in the low care condition (LCLS  $M=2.95$ ,  $SD=1.5$ ; LCHS  $M=3.09$ ,  $SD=1.7$ ) were significantly lower than in the high care conditions (HCHS  $M=5.60$ ,  $SD=1.5$ ; HCLS  $M=5.47$ ,  $SD=1.6$ ;  $F(3, 383) = 82.770$ ,  $p < .001$ ). Direct comparisons between mean ratings of care showed that ratings of care in the low care conditions did not differ significantly from each other ( $p = 1$ ), and ratings of care in the high care conditions did not differ significantly from one another ( $p = 1$ ). Ratings of care in the HCLS and HCHS conditions were significantly higher than ratings of care in the LCHS and LCLS conditions ( $p < .001$  in both cases). Additionally, participants gave similar answers to one another across different conditions as to how common the kind of mistake depicted in the vignette is (LCLS  $M=3.92$ ,  $SD=1.6$ ; LCHS  $M=4.17$ ,  $SD=1.6$ ; HCLS  $M=3.77$ ,  $SD=1.3$ ; HCHS  $M=4.09$ ,  $SD=1.5$ ). The mean ratings of these assessments did not differ significantly ( $F(3, 383) = 1.358$ ,  $p = .26$ ).

#### 4.3 Discussion

Study 3 mirrored the structure of Studies 1 and 2, but it employed a different vignette, in order to verify that the findings from the previous two studies were not specific to the vignettes used in those studies. Replicating the findings of Studies 1 & 2, in Study 3

participants judged that Randy is more at fault, that Randy's boss is more warranted to blame him, that he should feel more guilty, and that he is more morally responsible for not submitting the form when his stress levels are low than if his stress levels are high. As for the effects of care, the results were mixed. While Study 3 did not reveal an effect of care for fault and blame, it did replicate the effects of guilt and moral responsibility. Thus, as in Study 2, participants judged that Randy should feel more guilt and that Randy is more morally responsible for not submitting the form when he does not care than when he does care. In sum, the present study replicates the effects of stress that we found in Studies 1 & 2, but does not replicate all the effects of care that we found in Studies 1 & 2. Because the vignettes were changed in Study 3, we did not include a post-hoc check for effect of Study (as we did in sec. 3.3).

## 5. General Discussion

We posited four hypotheses about responsibility for forgetfulness in the Introduction:

- (H1) People will judge that agents are sometimes responsible for their forgetfulness.
- (H2) These judgments will be sensitive to the level of stress that the agent experiences while being forgetful.
- (H3) Stress will provide a partial excuse, but it will not function as a total excuse for some of the consequences of being forgetful.
- (H4) Judgments of responsibility will not be sensitive to how much the agent cares about performing the forgotten action.

The current studies lend support to hypotheses (H1)-(H3). The average ratings of fault, blame, and guilt in Study 1 were all above the midpoint. In Studies 2 & 3, participants on average thought that the target was at least somewhat responsible for being forgetful. We think that this indicates that participants were disposed to judge that agents are sometimes responsible for forgetfulness in accordance with hypothesis (H1).

In all three studies, judgments of responsibility were not maximally high. This fits with expectations given previous studies on the asymmetry between judgments of responsibility for deliberate and mistaken wrongdoing (Cushman 2008; Decety, Michalska, and Kinzler 2012; Killen et al. 2011; Young et al. 2007; Cushman et al. 2009). There is some similarity, then, between judgments of responsibility for mistaken wrongdoing and responsibility for forgetfulness in that both conditions result in judgments of reduced responsibility (relative to deliberate wrongdoing). However, the reason for the reduction might be different and future work could focus on potential differences between the two conditions.

Our second hypothesis (H2) was that the level of stress would affect judgments of responsibility. The results from Studies 1, 2, and 3 lend support to (H2), as it was found that stress levels had significant effects on judgments of blame, fault, guilt, and moral responsibility. These effects indicate that perceived level of stress modulates judgments of responsibility.

Our results support a different claim, as well, namely that stress partially excuses forgetting when the agent is not at fault for the stress. This would explain the change in judgments in the stress conditions between Study 1 and Studies 2 & 3. If, as we suspect, participants judge that Randy is at fault for being stressed in Study 1 (because he engaged in poor financial planning), then only non-culpable stress affects folk assessments of responsibility for forgetting. These results provide good (but not decisive) evidence for hypothesis (H3) that stress partially excuses responsibility for forgetting while care does not.

Importantly, stress in cases of forgetting does not function as a total excuse. A total excuse is some condition or state the presence of which provides total mitigation (cf.

Strawson 2003: 78). Stress does not function as a total excuse, in this sense, because people continued to judge that the target was somewhat responsible for his mistake even in the high stress scenarios. Stress functions as a partial excuse, where a partial excuse is some condition or state the presence of which provides partial mitigation for wrongdoing (Strawson 2003: 77; cf. Watson 1987). This provides more evidence for (H3).

Our fourth hypothesis (H4) claimed that level of care would not affect judgments of responsibility. As both variables had at least some significant effects on judgments of responsibility, our studies do not confirm (H4). However, three findings suggest that (H4) is likely true. First, the effect sizes of care for all dependent variables in Studies 1 and 2 were small, which provides some evidence that care may play only a minimal role in moral assessment. Second, we did not find significant effects of care on judgments of fault and blame in Study 3 (and minimal significant effects of care on judgments of guilt and moral responsibility), suggesting that care plays a minimal role in moral assessment. Moreover, the fact that we did not find certain effects of care in Study 3 suggests that these small effects might be specific to certain situations rather than generalizable across conditions. Finally, lack of change in effect sizes of care for all dependent variables between Studies 1 and 2 provides more evidence that care may play a minimal role in moral assessment. That is, even when extra information relating to care was provided in vignettes for Study 2, care ratings did not explain any more variance in the four outcomes assessed (see post-hoc test in Section 3.3).

### 5.1 The Nature of Responsible Agency

The results of our studies also contribute to the philosophical discussion mentioned in the introduction. In particular, our results show that we are disposed to

hold people responsible for (some of) their forgetfulness. This poses some problems for the valuational theory.

Recall that the valuational theory predicts that agents are not responsible for forgetfulness. In our studies, people judge that Randy is responsible for his forgetfulness despite the fact that he displays a high degree of care toward successfully getting the ingredients. This represents a divergence between the predictions of the valuational theory and folk practices of assigning responsibility.

In light of this, some valuationists (e.g., Doris 2015) try to dismiss these cases of forgetfulness as too rare to drive theory construction or theory revision (as noted in the introduction). Our studies, however, show that people think that these cases are somewhat common. Judgments about the frequency of these mistakes indicate that people believe that morally significant forgetfulness is not isolated. Mean ratings of frequency are above (or very near) the midpoint in both studies (Study 1: HCHS  $M=5.26$ ; LCHS  $M=5.68$ ; HCLS  $M=5.46$ ; LCLS  $M=5.32$  / Study 2: HCHS  $M=4.65$ ; LCHS  $M=5.01$ ; HCLS  $M=4.95$ ; LCLS  $M=4.71$  / Study 3: LCLS  $M=3.92$ ,  $SD=1.6$ ; LCHS  $M=4.17$ ,  $SD=1.6$ ; HCLS  $M=3.77$ ,  $SD=1.3$ ; HCHS  $M=4.09$ ,  $SD=1.5$ ), suggesting that valuational theorists cannot simply dismiss cases of responsibility for forgetfulness as anomalies or outliers that do not require explanation.

Finally, valuational theories often consider the pretheoretical credibility of the view as a reason to adopt the theory. Given our results, we think that the valuationist can no longer claim, unequivocally, that pretheoretical credibility favors the valuational theory. Insofar as there is significant divergence between our folk practices and the predictions of the valuational theory, this should be considered a significant cost of the theory. A proponent of the valuational theory might note that the frequency ratings in

Study 3 are much lower than those in Studies 1 & 2. We do not think this finding provides support for the valuational theory. One thing that might explain this is that submitting an online reimbursement request, if it is part of one's job, is likely not a common part of one's job. This might explain why people rated the mistake as less frequent than forgetting to stop at the store on the way home from work (a, presumably, much more common occurrence).

In addition to these reasons for doubting the valuational theory, our studies provide some positive evidence for the capacitarian theory. The fact that stress explains a significantly greater degree of variability in judgments of fault, blame, and guilt, aligns with the predictions of the capacitarian theory. In Study 3, there was no significant main effect of care on either judgments of fault or judgments of blame. This, we think, is more evidence in favor of the capacitarian theory. However, level of care did have some effect on judgments of fault, guilt, blame, and moral responsibility even if these effects were minimal or less explanatorily relevant than stress. This, again, is why our data does not confirm (H4).

At this point, a valuational theorist might object. After all, the things that cause us stress (especially high degrees of stress) are indicative of our cares. So, when Randy is under high stress, it is really the case that he is focused on *other* things that he cares about (like his job). Randy's stress, then, is really an indication of a deeper network of cares. And when people assess Randy's behavior under conditions of high stress, they are taking that stress as a proxy for his cares. So the valuational theorist can explain our data within her framework.

Of course, it is open to the valuational theorist to interpret the data in this way. We cannot determine, based on these studies, what people think of Randy's stress.

However, if it were the case that people were using stress as a proxy for Randy's cares (so that judgments of responsibility were really just taking Randy's cares into account), then variations in care across conditions should have explained more of the variance in responses than we observed. As our post-hoc testing revealed (sec. 3.3), the Care variable explains a negligible amount of the variance in responsibility judgments across conditions. Stress explains *almost all* of the variance. If, however, the valuationist is correct in her objection, then the effect of the Care variable itself should be more significant. As it is, our results show that people did not treat Care and Stress equivalently or even similarly. Even further, Care did not even have a significant effect on judgments of fault and blame in Study 3. If the Care is as important to responsible agency as valuationists suppose, one should not expect to find a lack of significant effect of care.

At this point, the valuational theorist has a choice. Either she can explain why, even though stress is a proxy for care, the Care variable ends up doing almost no explanatory work in explaining variations in judgments across conditions. Or, she can propose the valuational theory as a revisionary conception of responsible agency. This latter option releases the valuational theorist from the task of fitting her theory to the data about folk judgments. But the cost is that there is little reason to adopt the theory (especially since there is a plausible alternative view of responsible agency that fits with the data about folk judgments).

## 5.2 Future Studies

There are several questions that these studies raise for future research. The first is simply whether these judgments correlate with judgments of responsibility for other kinds

of mistakes beyond forgetting. Future studies could compare assessments of responsibility for different kinds of mistakes.

Second, future studies could also compare judgments of responsibility for consequences that result from forgetful omissions to consequences that result from forgetful actions. That is, Jane might forget to give a message to John, or Jane might forget not to tell a secret to John, both of which result in something for which Jane is responsible. Are the conditions that affect judgments of responsibility the same in both cases?

Third, what other kinds of mitigating factors are there? Stress and care are certainly not the only elements in our folk categories of excuse and mitigation. Additionally, are there any mitigating conditions that produce a more dramatic effect on assessments of responsibility than stress? For example, does the amount of time that passes between forming a plan and needing to execute that plan affect judgments of responsibility for forgetting to execute the plan? Or would there be diminished judgments of responsibility if an external device (like an iPhone) failed to send a reminder that one set *and* one forgot to execute some plan (or forgot to perform some promised action, etc.)?

Fourth, do different kinds of stress have differential effects on assessments of responsibility for forgetfulness? That is, are people's assessments of responsibility in these cases sensitive to fine-grained differences between kinds of stress? For example, does it make a difference if stress is triggered by thoughts about oneself or by thoughts about others (like those one cares about)? Does it make a difference if stress is internally or externally triggered? Do different kinds of stress have differential effects on reasons-responsiveness? Future studies might test the effects of these different kinds of stress.

Fifth, what do people intuitively take the scope of stress to be? For this paper, we assumed that stress is a species of distraction, where stress is some psychological state of an agent caused by an unexpected present event or unpredictable future event that is connected to dispositions to respond to that event on the part of the agent. On this conception of stress, stress could be positive or negative and doesn't depend on any distinctive affective response from the agent. However, the folk conception of stress might not map exactly onto this theoretical construct. Perhaps the folk conception of stress encompasses only negative stress. Alternatively, the folk conception of stress might include some affective component as essential to stress. Study 2 replicated the effect of Study 1 despite the fact that in Study 1 the Low Stress condition was positive stress. And, the effect of stress in Study 2 (where the Low Stress condition was negative stress) was *larger* than Study 1. This suggests that the folk take positive stress, *ceteris paribus*, as a mitigating factor, though this could only be confirmed in additional studies with more targeted measures. In all, the psychology of stress and stress attribution remains underexplored. Additionally, future *conceptual* work might focus on generating a fully general definition of the psychological state of stress that also specifies the scope of stress.

Finally, what is the relationship between folk judgments of fault, blame, and guilt? In Studies 2 & 3, judgments of moral responsibility did not line up closely with any other categories of judgments. Further studies could investigate the similarities and differences between these concepts and the conditions under which people apply these concepts in various circumstances. Additionally, future studies might try to determine what the relations are among the concepts of fault, blame, guilt, and moral responsibility.

## **References**

- Ahrens, M. (2016). *Home fires involving cooking equipment* (Quincy, MA: NFPA Fire Analysis and Research).
- Amaya, S. (2013). "Slips," *Noûs* 47:3, 559-76.
- Amaya, S. and Doris, J. (2014) "No Excuses: Performance Mistakes in Morality," In J. Clausen and N. Levy (eds.) *Handbook of Neuroethics*, 253-272.
- Arnsten, A.F.T. (2015). "Stress weakens prefrontal networks: molecular insults to higher cognition," *Nature Neuroscience* 18:10, 1376-1385.
- Arpaly, N. and Schroeder, T. (2014). *In Praise of Desire* (Oxford: Oxford University Press).
- Björnsson, G. (2017). "Explaining Away Epistemic Skepticism about Culpability," In *Oxford Studies in Agency and Responsibility*, vol. 4, ed. D. Shoemaker (Oxford: Oxford University Press), 141-64.
- Clarke, R. (2014). *Omissions: Responsibility, Agency, and Metaphysics* (Oxford: Oxford University Press).
- Clarke, R. (2016). "Ignorance, Revision, and Commonsense," In P. Robichaud, J.W. Wieland (eds.) *Responsibility: The Epistemic Condition* (Oxford: Oxford University Press).
- Collins, J. (2006). "Crime and Parenthood: The Uneasy Case for Prosecuting Negligent Parents," *Northwestern University Law Review* 100:2, 807-855.
- Cushman, F. (2008). "Crime and Punishment: Distinguishing the roles of causal and intentional analyses in moral judgment," *Cognition* 108, 353-380.
- Cushman, F., A. Dreber, Y. Wang, and J. Costa. (2009). "Accidental outcomes guide punishment in a 'trembling hand' game," *PloS One* 4:8, e6699.
- Darwall, S. (2002). *Welfare and Rational Care* (Princeton: Princeton University Press).
- Decety, J. and S. Cacioppo. (2012). "The speed of morality: A high-density electrical neuroimaging study," *Journal of Neurophysiology* 108, 3068-3072.
- De Freitas, J., Sarkissian, H., Newman, G.E., Grossman, I., De Brigard, F., Luco, A., and Knobe, J. (2017). Consistent Belief in a Good True Self in Misanthropes and Three Independent Cultures. *Cognitive Science*. doi: 10.1111/cogs.12505.
- Finkel, N.J. and J.L. Groscup. (1997). "When Mistakes Happen: Commonsense Rules of Culpability," *Psychology, Public Policy, and Law* 3:1, 65-125.
- Fischer, J.M. and Ravizza, M. (1998). *Responsibility and Control* (Cambridge: Cambridge University Press).

- Frankfurt, H.G. (1983). "What we are morally responsible for," In *How Many Questions? Essays in Honor of Sidney Morgenbesser*, ed. L. Cauman, I. Levi, C. Parsons, R. Schwartz (Indianapolis, IN: Hackett Publishing), 321-335.
- Frankfurt, H.G. (1999). "Autonomy, Necessity, and Love," In *Necessity, Volition, and Love* (New York: Cambridge University Press), 129-41.
- Goldfarb, E.V., Froböse, M.I., Cools, R., and Phelps, E.A. (2017). "Stress and cognitive flexibility: cortisol increases are associated with enhanced updating and impaired switching," *Journal of Cognitive Neuroscience* 29:1, 14-24.
- Guerrero, A.A. (2007). "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution," *Philosophical Studies* 136, 59-97.
- Harman, E. (2011). "Does Moral Ignorance Exculpate?" *Ratio* 24, 443-468.
- Hart, H.L.A. (1968). *Punishment and Responsibility* (Oxford: Oxford University Press).
- Hieronymi, P. (2008). "Responsibility for Believing," *Synthese* 161:3, 357-73.
- Hollon, N.G., Burgeno, L.M., Phillips, P.E.M. (2015). "Stress effects on the neural substrates of motivated behavior," *Nature Neuroscience* 18:10, 1405-12.
- Hurd, H. and M. Moore (2011). "Punishing the Awkward, the Stupid, the Weak, and the Selfish: The Culpability of Negligence," *Criminal Law and Philosophy* 5:2, 147-198.
- Husak, D. (2011). "Negligence, Belief, Blame and Criminal Liability: The Special Case of Forgetting," *Criminal Law and Philosophy* 5:2, 199-218.
- Jaworska, A. (1999). "Respecting the Margins of Agency: Alzheimer's patients and the capacity to value," *Philosophy and Public Affairs* 28, 105-38.
- Keegan, H. (2016). "The 8 most costly cooking mistakes in Springfield—and how they could've been avoided," *Springfield News-Leader*, published online 12/19/2016. From <http://www.news-leader.com/story/news/local/ozarks/2016/12/19/8-most-costly-cooking-mistakes-springfield-and-how-they-couldve-been-avoided/94730926/>.
- Killen, M., K.L. Mulvey, C. Richardson, N. Jampol, and A. Woodward. (2011). "The accidental transgressor: Morally-relevant theory of mind," *Cognition* 119, 197-215.
- Levy, N. (2009). "Culpable Ignorance and Moral Responsibility," *Ethics* 119, 729-741.
- Martin, J.W. and F. Cushman. (2016). "Why we forgive what can't be controlled," *Cognition* 147, 133-143.

- Maxwell, B. and Racine, E. (2012). "Does Neuroscience Research on Early Stress Justify Responsive Childcare? Examining Interwoven Epistemological and Ethical Challenges," *Neuroethics* 5:2, 159-72.
- Mele, A. (2006). "Practical Mistakes and Intentional Actions," *American Philosophical Quarterly* 43:3, 249-260.
- Malle, B.F., Guglielmo, S., and Monroe, A.E. (2012). "A Theory of Blame," *Psychological Inquiry* 25:2, 147-186.
- Mason, E. (2015). "Moral ignorance and blameworthiness," *Philosophical Studies* 172, 3037-3057.
- McEwen, B.S., Bowles, N.P., Gray, J.D., Hill, M.N., Hunter, R.G., Karatsoreos, I.N., and Nasca, C. (2015). "Mechanisms of stress in the brain," *Nature Neuroscience* 18:10, 1353-63.
- McGeer, V. and Pettit, P. (2015). "The Hard Problem of Responsibility," In *Oxford Studies in Agency and Responsibility*, vol. 3, ed. D. Shoemaker (Oxford: Oxford University Press), 160-88.
- Murray, S. (2017). "Responsibility and Vigilance," *Philosophical Studies* 174:2, 507-27.
- Nelkin, D.K. (2016). "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness," *Noûs* 50:2, 356-378.
- Pelletiere, N. (2016). "'Forgotten Baby Syndrome': A Parent's Nightmare of Hot Car Death," *ABC News*, published online 07/14/2016. From <http://abcnews.go.com/Lifestyle/forgotten-baby-syndrome-parents-nightmare-hot-car-death/story?id=40431117>.
- Raz, J. (2011). *From Normativity to Responsibility* (Oxford: Oxford University Press).
- Reynolds, S.J., Owens, B.P, and Rubenstein, A.L. (2012). "Moral Stress: Considering the Nature and Effects of Managerial Moral Uncertainty," *Journal of Business Ethics* 106:4, 491-502.
- Rosen, G. (2015). "The Alethic Conception of Moral Responsibility," In R. Clarke, M. McKenna, and A. Smith (eds.) *New Essays on Moral Responsibility* (Oxford: Oxford University Press).
- Sapolsky, R.M. (2015). "Stress and the brain: Individual variability and the inverted-U," *Nature Neuroscience* 18:10, 1344-46.
- Scanlon, T.M. (2008). *Moral Dimensions* (Cambridge, MA: Belknap Press).

- Scult, M.A. (2017). "Flexible Adaptation of Brain Networks during Stress," *Journal of Neuroscience* 37:15, 3992-3994.
- Shoemaker, D. (2011). "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility," *Ethics* 121:3, 602-632.
- Shoemaker, D. (2015). *Responsibility from the Margins* (Oxford: Oxford University Press).
- Smith, A.M. (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115:2, 236-271.
- Smith, A.M. (2008). "Control, Responsibility, and Moral Assessment," *Philosophical Studies* 138:3, 367-392.
- Smith, H.M. (1983). "Culpable Ignorance," *Philosophical Review* 94:4, 543-571.
- Sripada, C. (2015). "Moral Responsibility, Reasons, and the Self," In D. Shoemaker (ed.) *Oxford Studies in Agency and Responsibility*, vol. 3 (Oxford: Oxford University Press), 242-64.
- Sripada, C. (2016). "Self-expression: a deep self theory of moral responsibility," *Philosophical Studies* 173, 1203-1232.
- Strawson, P.F. (2003). "Freedom and Resentment," In G. Watson (ed.) *Free Will* (Oxford: Oxford University Press), 72-93.
- van Inwagen, P. (1978). "Ability and Responsibility," *Philosophical Review* 87:2, 201-224.
- van Inwagen, P. (2008). "How to Think about the Problem of Free Will," *Journal of Ethics* 12, 327-341.
- van Inwagen, P. (2017). "The Problem of Fr\*\* W\*ll," In *Thinking About Free Will* (Cambridge: Cambridge University Press), pp. 192-209.
- Vargas, M. (2013). *Building Better Beings* (Oxford: Oxford University Press).
- Watson, G. (1987). "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme," In *Responsibility, Character and the Emotions*, ed. F.D. Schoeman (Cambridge: Cambridge University Press), 256-286.
- Watson, G. (1996). "Two Faces of Responsibility," *Philosophical Issues* 24:2, 227-48.
- Weingarten, G. (2009). "Fatal Distraction," *The Washington Post*, 03/08/09. From <http://www.washingtonpost.com/wpdyn/content/article/2009/02/27/AR2009022701549.html>

Woolfolk, R.L., Doris, J.M., and Darley, J.M. (2006). Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility,” *Cognition* 100, 283-291.

Yaffe, G. (2012). “Intoxication, Recklessness and Negligence,” in *Ohio State Journal of Criminal Law* 9, 545-583.

Young, L., F. Cushman, M. Hauser, and R. Saxe. (2007). “The neural basis of the interaction between theory of mind and moral judgment,” *Proceedings of the National Academy of Sciences of the United States of America* 104, 8325-40.

Young, L., S. Nichols, and R. Saxe (2010). “Investigating the neural and cognitive basis of moral luck: It’s not what you do but what you know,” *Review of Philosophy and Psychology* 1, 333-349.

Young, L. and L. Tsoi. (2013). “When Mental States Matter, When They Don’t, and What That Means for Morality,” *Social and Personality Psychology Compass* 7/8, 585-604.

Zimmerman, M. (2008). *Living with Uncertainty* (Cambridge: Cambridge University Press).