

جامعة نيويورك أبوظبي

 NYU | ABU DHABI



# **Surveillance cues do not enhance altruistic behavior among anonymous strangers in the field**

**Erik J. Koornneef, Aurelie Dariel, Iffat Elbarazi,  
Ahmed R. Alsuwaidi, Paul B.M. Robben and Nikos Nikiforakis**

**Working Paper # 0017**

**May 2018**

**Division of Social Science Working Paper Series**

**New York University Abu Dhabi, Saadiyat Island P.O Box 129188, Abu Dhabi, UAE**

**<https://nyuad.nyu.edu/en/academics/divisions/social-science.html>**

# Surveillance cues do not enhance altruistic behavior among anonymous strangers in the field

Erik J. Koornneef<sup>\*</sup>, Aurelie Dariel<sup>†</sup>, Iffat Elbarazi<sup>‡</sup>  
Ahmed R. Alsuwaidi<sup>‡</sup>, Paul B.M. Robben<sup>\*</sup> and Nikos Nikiforakis<sup>†§</sup>

## Abstract

The degree of altruistic behavior among strangers is an evolutionary puzzle. A prominent explanation is the evolutionary legacy hypothesis according to which an evolved reciprocity-based psychology affects behavior even when reciprocity is impossible, i.e., altruistic behavior in such instances is maladaptive. Empirical support for this explanation comes from laboratory experiments showing that surveillance cues, e.g., photographs of watching eyes, increase altruistic behavior. A competing interpretation for this evidence, however, is that the cues signal the experimenter's expectations and participants, aware of being monitored, intentionally behave more altruistically to boost their reputation. Here we report the first results from a field experiment on the topic in which participants are unaware they are being monitored and reciprocity is precluded. The experiment investigates the impact of surveillance cues on a textbook example of altruistic behavior – hand hygiene *prior* to treating a 'patient'. We find no evidence surveillance cues affect hand hygiene, despite using different measures of hand-hygiene quality and cues that have been previously shown to be effective. We argue that surveillance cues may have an effect only when participants have reasons to believe they are actually monitored. Thus they cannot support claims altruistic behavior between strangers is maladaptive.

*Keywords:* altruism, watching eyes, strangers, natural field experiment

---

<sup>\*</sup> Institute for Health Policy and Management, Erasmus University, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands.

<sup>†</sup> Division of Social Science, New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, UAE.

<sup>‡</sup> Institute of Public Health, UAE University, P.O. Box 17666, Al Ain, UAE. <sup>‡</sup> Department of Pediatrics, UAE University, P.O. Box 17666, Al Ain, UAE.

<sup>§</sup> Corresponding author: [nikos.nikiforakis@nyu.edu](mailto:nikos.nikiforakis@nyu.edu)

## **Introduction**

The degree of altruistic behavior among strangers in modern societies is a major evolutionary puzzle [1,2]. A prominent explanation is the ‘evolutionary legacy hypothesis.’ It posits that the human brain evolved in ancestral conditions that differed radically from those in modern environments [3,4]. Although nowadays many encounters are with anonymous strangers, for much of our evolutionary past, humans interacted repeatedly in small social groups where one’s reputation was constantly at stake, leading to the evolution of cognitive mechanisms to automatically identify reputation-building opportunities [5-7]. According to the evolutionary legacy hypothesis, individuals may behave altruistically in anonymous one-shot interactions due to an uncontrolled, automatic reaction aimed to bolster one’s good reputation in anticipation of positive reciprocity, even when such opportunities do not exist; i.e., the observed altruistic behavior between anonymous strangers is maladaptive [8,9].

Empirical support for this hypothesis comes from laboratory experiments showing that reputation-related surveillance cues such as displaying photographs of watching eyes promote altruistic behavior, i.e., actions which benefit another individual at a personal cost [10-12]. Since the cues do not affect one’s reputation, participants are anonymous to each other and direct reciprocity is precluded by design, the effect has been attributed to the automatic activation of one’s reciprocity-based psychology [8,9]. A problem with this interpretation, however, is that participants are not anonymous to the experimenter who may be observing their choices or can easily infer them from their earnings. A competing interpretation therefore is that the increase in altruism is due to an ‘experimenter effect’ [13-15]: surveillance cues signal the experimenter’s expectations to participants who intentionally react to the stimulus in a way they believe would boost their reputation with him/her. In other words, even though direct reciprocity between participants is

precluded, indirect reciprocity concerns may still play a role. In line with this interpretation, survey evidence shows the effect on altruistic behavior is mediated by participants' expectation of reward by "a third party who was monitoring them" [16].

One way to eliminate the possibility of an experimenter effect is to conduct field experiments such that individuals are unaware they are participating in a study. In order to provide clear evidence that altruistic behavior is maladaptive, however, certain conditions need to be satisfied such that alternative explanations are ruled out. In particular, it is critical that all encounters in the experiment are anonymous, one-shot and reciprocity of any kind is precluded. It is also desirable that exposure to the cues is brief as habituation with the false stimulus may attenuate the effect through intentional brain processes [7,12]. If surveillance cues are found to increase altruistic behavior in such circumstances this would support the hypothesis that altruistic behavior between anonymous strangers is maladaptive. If surveillance cues have no effect, it would suggest that previous findings may have been due to an *intentional* decision taken to bolster one's reputation with the experimenter. Here, we present the first evidence from such a field experiment.

We take advantage of a unique opportunity to study altruistic behavior in a setup which meets all the aforementioned requirements. This distinguishes our experiment from previous field studies in which surveillance cues were displayed in *public* spaces over an *extended* period of time [17-25]. As we explain in the last section of our paper, these studies were designed to address different research questions. As such, the positive effect of surveillance cues on altruistic behavior in these studies suggests a potentially useful, low-cost, policy intervention, but it cannot support the claim that altruistic behavior among strangers is maladaptive as many of the aforementioned conditions are not satisfied and indirect reciprocity opportunities exist.

In our experiment we investigate how two distinct surveillance cues impact the quality of hand hygiene by medical students before treating a ‘patient.’ Hand hygiene (HH) is a general term used to describe the process of removing microorganisms with a disinfectant agent such as alcohol, or soap and water [26]. Appropriate HH among healthcare workers is considered by some to be the most effective measure to prevent healthcare-associated infections [27], which are associated with 50,000 and 99,000 deaths each year in Europe and the USA, respectively [28], and annual hospital costs between \$28.4 to \$33.8 billion USD [26]. It is estimated that a one-percent improvement in the quality of hand hygiene could save approximately \$40,000 USD per year in a 200-bed hospital for a single type of infection [29]. That is, *how* one washes his or her hands is critical. Accordingly, compliance with HH guidelines has been identified by the World Health Organization (WHO) as a first priority in health-care facilities [28].

Appropriate HH *prior* to treating a patient is a textbook example of altruistic behavior. According to WHO’s guidelines, when contact with a patient is not invasive – as is the case in our experiment – a healthcare provider must follow a specific technique to thoroughly wash his/her hands both before and after contact, for 40 to 60 seconds each time (when contact is invasive, the duration should be between 120 to 300 seconds) [28]. HH prior to treating a patient is not only costly, taking time and effort, but also it does not benefit the healthcare provider directly, only the patient whose chances of a healthcare associated infection are estimated to decrease between 15 and 30% [30]. Both the cost to the practitioner [31] and the lack of individual benefits [32] from HH prior to treating a patient have been cited as prime reasons for why compliance with WHO’s guidelines is low. In support of the idea that HH prior to treating a patient constitutes an altruistic act is the evidence that compliance with HH guidelines is substantially higher *after* contact with the patient [31,33]. In the concluding section, we present results from a survey showing that concerns for the

welfare of the patient indeed appear to be the primary reason for washing hands prior to treating a patient in our experiment.

### **The experiment**

The experiment was conducted in a large university, which is well-regarded locally for its medical program: the United Arab Emirates University (see section S1 in Supplementary Information, *SI*). Participants were advanced undergraduate students in the Doctor of Medicine (MD) program who had completed training modules in the basic principles of clinical practice, including infection prevention and HH in accordance with WHO guidelines. For the experiment, we took advantage of a unique opportunity offered by the program for students to *privately* practice their clinical skills – a *Practice* Objective Structured Clinical Examination (POSCE). The *official* OSCE is a critical part of all MD programs aimed to formally evaluate one’s clinical competence. Medical students in the OSCE are observed and evaluated by faculty members as they go through a series of stations, interviewing, examining and treating different standardized patients who present some type of medical problem. The POSCE was identical to the OSCE, with two crucial differences: (i) faculty members were not present to observe or evaluate the competence of the students, and (ii) participants remained anonymous throughout the process. Students were fully aware that the purpose of the exercise was for them to practice their skills without being judged or evaluated.

A note outside the ‘patient’s room’ informed students that their main task was to take the blood pressure of a standardized patient (see section S4 in *SI*). Medical students are aware that, whenever they are having physical contact with a standardized patient, there is a real risk of contaminating him/her. Participants therefore know that best clinical practice requires they wash their hands carefully immediately prior to measuring the standardized patient’s blood pressure, following the WHO’s HH guidelines. At the same time, participants are not monitored and, like with the OSCE,

each practice slot lasts ten minutes – this is signified by automated bells in the corridors, which were meant to reinforce the fact that the POSCE was not monitored. During this time, they had to briefly interview the standardized patient, wash their hands, measure blood pressure, wash their hands again, and provide feedback to the patient. Given their limited experience with the blood measurement instruments, participants had an incentive to take advantage of this one-off opportunity and spend most of their time practicing measuring blood pressure as it is likely to be relevant in the official OSCE. Therefore, there is a non-trivial cost for participants from properly washing their hands, but incurring the cost benefits the standardized patient. It should be noted that students could not benefit patients by expediting HH as each POSCE slot lasted *exactly* ten minutes, i.e., they could leave neither earlier nor later.

The experimental treatments varied the surveillance cues which were displayed, approximately at eye-level, above a wash basin (see Fig.1, and section S7 in *SI*) and underneath the standard HH poster by WHO explaining in detail appropriate HH (see section S8 in *SI*). Due to the HH guidelines and the limited time of the session, exposure to the surveillance cue was necessarily brief (<60 seconds). Participants were randomly assigned to treatments/cues (see section S1 in *SI*). The *Baseline* condition, like previous studies, consists of a non-reputation-related image – the picture of a tree. In the *Eyes* treatment, a pair of stern-looking male eyes was displayed. This particular image was chosen as it has been previously associated with a large positive effect on HH, i.e., a 122% increase relative to a baseline condition when the cue was placed in a public space over an extended time period [25]. This was important as it was uncertain *ex ante* how large a sample we could hope to attract. Ultimately, the turnout was substantial and higher than we had expected: 114 students out of an eligible student population of 330. With this sample size, our tests have sufficient power to detect treatment differences substantially smaller than those in King et al.

[25] (see section S3 in *SI*). Note that in order to be exposed to the treatment manipulation, participants had to go to the wash basin to wash their hands. Some participants in our sample had to be prompted to do so by the standardized patient. Our results are unaffected if we exclude these participants from the analysis (see section S2 in *SI*).

**Figure 1** – Picture of the wash basin in the private examination room featuring the watching eyes



The *Camera* treatment is the first of its kind in the literature. In this treatment, the picture of a CCTV camera was placed over the wash basin. Such CCTV cameras are omnipresent in the country of our study, although none was available in the examination room. If people have developed cognitive mechanisms to *automatically* identify reputation-building opportunities through millennia, and this is the cause for the changes in altruistic behavior when reputation-related cues are presented, then we would expect to observe an effect in the Eyes but not in the



Camera treatment [8]. A difference in HH between our Camera and Baseline conditions could be interpreted as evidence participants are concerned about actually being watched.

Apart from informing students they would need to measure blood pressure, the briefing note outside the room explained that the only people inside the room will be two simulated patients that will take turns being the ‘patient’. (Rooms were spacious: approx. 25 square meters or 270 square feet, see section S6 in *SI*.) This is standard practice as a person should not have his/her blood pressure measured repeatedly. The simulated patients (RAs) were trained to appear indifferent to the actions of the medical students. At any given point, one RA waited for his/her blood pressure to be measured, and the other –seated at a faraway corner of the room – waited for his turn, while filling out a Sudoku book. In actuality, this individual was covertly monitoring the student’s HH practice (see sections S5 in *SI*). All RAs had been professionally trained on how to evaluate the quality of one’s HH. Crucially, the simulated patients were selected such that they were completely unknown to the students (with one exception, see section S2 in *SI*).

Like in previous studies, direct reciprocity is prevented by design: not only were standardized patients in a passive role, but they were also trained to appear bored and indifferent to the POSCE. To preclude indirect reciprocity, encounters had to be anonymous such that reputational concerns could not affect altruistic behavior. For this reason, students and simulated patients were explicitly instructed not to share their identities. As both ‘patients’ would remain in the room at the end of the session to receive the next medical student, while the practicing student would leave, it was clear to participants that encounters were one-shot and that there would be no opportunities for the ‘patients’ to reciprocate, either directly or indirectly. Further, to ensure observers were blind to our treatment manipulation, the RA who acted as the patient changed the poster before the next

participant entered the room so that the observer was not aware which poster was displayed at any point in time.

## Results

Our measure of altruistic behavior is the quality of hand hygiene prior to treating the patient. As mentioned, how a medical practitioner washes his/her hands is of critical importance for minimizing the risk of an infection. The survey evidence presented in the concluding section suggests that participants were well aware of this. In order to evaluate the quality of hand hygiene amongst participants, we use three distinct measures from the WHO guidelines about HH. First, we consider the time spent washing hands. Second, we study the quality of hand coverage, i.e., the extent to which a participant washed all surfaces of his/her hands. Third, we consider compliance with a rule prescribing participants use a tissue to switch off the tap, after finishing washing their hands. For simplicity and brevity, we present the means of these variables in the figures below; the distributions of these variables can be found in section S11 in *SI*.

Fig. 2 shows that, across treatments, participants on average washed their hands for slightly more than 20 seconds. While this may be more than the time spent by many adults washing their hands, it falls considerably short of the minimum recommended duration stated by WHO (40 seconds). This implies that there are good conditions for our treatments to increase the quality of hand hygiene and, in this instance in particular, the time spent washing. However, we find no statistically significant differences across treatments (Eyes vs. Baseline:  $P=0.69$ ,  $N=71$ ; Camera vs. Baseline:  $P=0.64$ ,  $N=79$ ; Mann-Whitney Test, two-tailed).

**Figure 2** –Average time spent washing hands across treatments (with 95-percent confidence intervals)

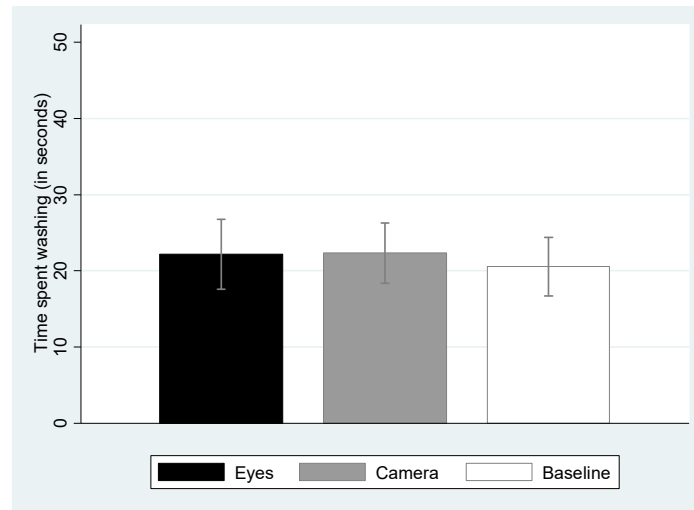


Fig. 3 presents the average quality of hand coverage across treatments. As mentioned, the RAs were professionally trained to evaluate the extent to which participants followed the WHO guidelines and, in this instance, covered adequately all hand surfaces. Performance was coded as 0 if the participant did not attempt to cover multiple surfaces (e.g., did a simple rub of the palms against each other), as 1 if the participant covered multiple but not all surfaces (e.g., did not wash thumbs), and 2 if the participant covered all surfaces. As can be seen in Fig. 3, average coverage is very similar in Eyes and Baseline, and statistically indistinguishable ( $P=0.99$ ,  $N=71$ ; Mann-Whitney Test, two-tailed). Although the quality of coverage is slightly higher in the Camera treatment, the difference with Baseline is statistically insignificant ( $P=0.18$ ,  $N=79$ ; Mann-Whitney Test, two-tailed).

**Figure 3** – Average quality of hand coverage across treatments (with 95-percent confidence intervals)

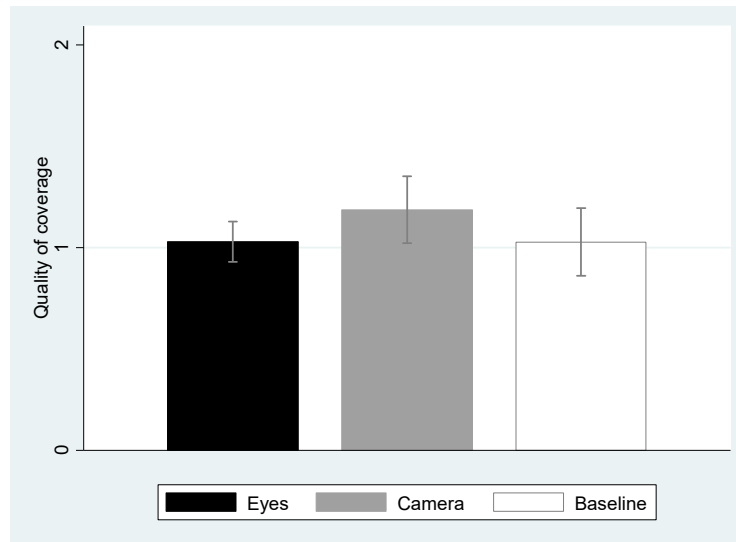
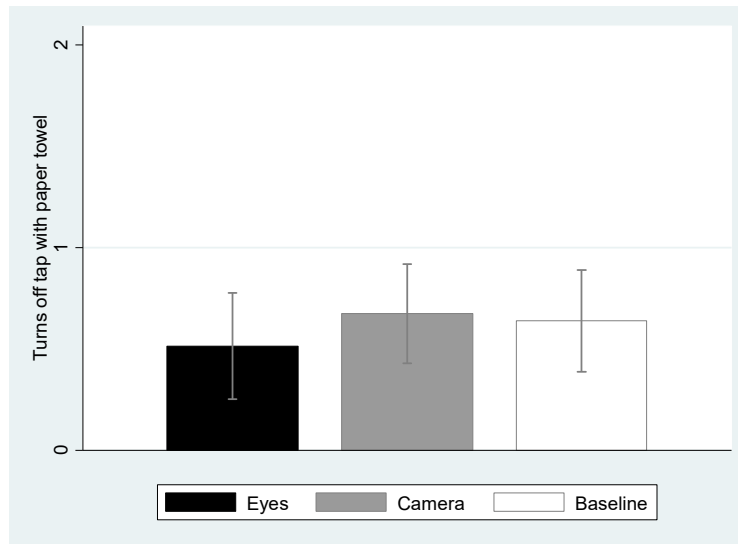


Fig. 4 presents the extent to which participants across treatments turned off the tap after washing their hands using a paper towel. This is critical in HH because a lot of bacteria can be found on the water tap. Participants have therefore been trained that not using a towel reduces considerably the efficacy of HH at combating disease transmission. Performance was coded as 0 if the participant did not use a paper towel at all, as 1 if the participant used a paper towel but improperly (e.g., used paper towel but also touched tap with bare hands), and 2 if the participant used properly a paper towel. Average compliance with this rule is lower in Eyes than in Baseline, although the difference is again statistically insignificant ( $P=0.39$ ,  $N=71$ ; Mann-Whitney Test, two-tailed). Compliance is similar in Camera and Baseline and statistically insignificant ( $P=0.90$ ,  $N=79$ ; Mann-Whitney Test, two-tailed). Neither the fraction of participants using a paper towel properly differs significantly across treatments (Eyes: 17.1%, Camera: 20.9%, Baseline: 16.7%; Eyes vs. Baseline:  $P=1.00$ ,  $N=71$ ; Camera vs. Baseline:  $P=0.78$ ,  $N=79$ ; Fisher Exact Test, two-tailed) nor the fraction of participants not using a towel at all (Eyes: 65.7%, Camera: 53.5%, Baseline: 52.7%; Eyes vs. Baseline:  $P=0.37$ ,  $N=71$ ; Camera vs. Baseline:  $P=1.00$ ,  $N=79$ ; Fisher Exact Test, two-tailed).

**Figure 4** – Average compliance with turning-off-tap-with-paper-towel rule across treatments (with 95-percent confidence intervals)



## Discussion

Our paper presents the first empirical test of the impact of surveillance cues on the altruistic behavior of anonymous strangers when reciprocity is precluded and participants are unaware they are being studied. These conditions are critical to obtain clear support for the evolutionary legacy hypothesis – a prominent explanation for altruistic behavior between strangers – according to which costly altruistic behavior in anonymous encounters is an anomaly, owing to our ancestral past and the development of automated, involuntary mechanisms for boosting one’s good reputation. Despite using cues that have been successfully used previously in the literature, we find no evidence surveillance cues increase the degree of altruistic behavior in our experiment. That is, our findings do not support the hypothesis that altruistic behavior among strangers is maladaptive.

One concern with all studies reporting null results is that this is due to the statistical tests being underpowered. This is clearly not the case in our experiment. Not only do we find no evidence across three distinct measures that the picture of eyes has a significant impact on altruistic behavior

in our experiment, but the effect itself is sometimes zero (Fig. 3) or negative (Fig. 4). By comparison, the effect of posting a picture of a camera over a wash basin – which as we argued could not be considered as supportive of the evolutionary legacy hypothesis – is also always insignificant and small in size, but at least it is always positive. Therefore, the overall lack of a significant effect cannot be attributed to insufficient statistical power.

Another concern may be that the lack of a positive effect is due to the fact that our experiment investigates the impact of surveillance cues on the quality of hand hygiene (intensive margin) but not on the decision to wash hands (extensive margin). Indeed, early evidence from dictator game experiments – in which an individual is assigned an amount of money and must decide how much of it to share with a passive recipient – suggested that surveillance cues may have a greater impact on the likelihood a ‘dictator’ shares a positive amount (extensive margin) than on the actual amount they share (intensive margin); combining margins the effect was often zero [11]. A recent meta-analysis of laboratory studies however contradicts these earlier results, finding no differential effect of cues on the extensive and the intensive margin [34]. Further, some field studies find the opposite result, i.e., that the impact of the cues is stronger on the intensive margin [19] or that a positive effect on charitable donations is obtained even when there are no differences in the proportion of donors responding to the cues [13]. Therefore, there exist neither clear empirical evidence nor theoretical reasons to expect the automatic activation of the reciprocity-based psychology will operate differently on the decisions on the extensive and intensive margins.

Although our design precludes both direct and indirect reciprocity by ensuring encounters are one-shot and all individuals involved (both participants and standardized patients) are unknown and anonymous to each other, we cannot rule out the possibility that, despite our efforts to avoid this, the presence of the standardized patients may have activated participants’ reciprocity-based

psychology already in the Baseline treatment, making it difficult to identify a treatment effect. However, it is worth emphasizing that similar concerns apply in laboratory environments. In fact, they are arguably greater: not only there are several participants in the lab *at the same time* – some of whom subjects may know personally – but their decisions are recorded by a computer and possibly observed by the experimenter. Even if this is not the case, participants – who often partake repeatedly in lab experiments – should anticipate that their final payment will ultimately reveal the extent of their altruistic behavior to the experimenter. If the reason for not observing a surveillance-cue effect in our experiment is the activation of the reciprocity-based psychology already in the Baseline condition, then it follows that the lab evidence on “watching eyes” cannot provide clear support either that altruistic behavior among strangers is maladaptive. It should also be noted that behavior across measures in our experiment falls considerably short of that described in the WHO guidelines. If participants were concerned about their reputation, one might have expected higher compliance with the guidelines than observed.

An altogether different concern with our study could be that hand hygiene prior to treating a patient is in fact *not* an altruistic act as we claimed. Although similar claims are common in the medical literature [31,33], one might wonder whether hand hygiene is regarded as altruistic, i.e., as conferring a benefit to the patient, in our particular context by medical students such those participating in our experiment. To address this concern, we administered a survey to 100 medical students with the same level of training and background as those who participated in our experiment (see section S9 in *SI*). Respondents were presented with a vignette designed to mimic the situation and incentives in our experiment, and asked whether they would wash their hands prior to treating the patient or not, and the reasons for their decisions. The survey also included a question to evaluate our claim that the quality of hand-hygiene matters when it comes to reducing

infection by asking participants whether they agree that washing hands for longer reduces the risk of infection for the patient.

Of the survey respondents who stated they would wash their hands prior to taking the blood pressure, 96.6% agreed with the statement that they would do so to avoid doing harm to the patient. We can reject the hypothesis that respondents neither agree nor disagree with the statement in favor of the alternative hypothesis that they agree with it ( $P < 0.01$ ,  $N = 87$ , Wilcoxon signed rank, two tailed). This supports our interpretation of hand hygiene prior to treating a patient as being an altruistic act as it is driven by a concern for the welfare of the patient. Participants were also more likely to agree with this reason for hand washing than with any of the other reasons ( $P < 0.01$  for all pairwise comparisons,  $N = 87$ , Wilcoxon signed rank, two tailed) indicating that the desire to do no harm to the patient is the main concern driving hand hygiene prior to contact with the patient. Perhaps unsurprisingly, we find that other concerns also appear to play a role (see sections S9 and S10 in *SI*), implying that hand hygiene prior to treatment is not driven exclusively by altruistic concerns for everyone. This, however, does not invalidate our analysis which only requires that altruistic motives are an important determinant of behavior in our experiment. We can also reject the hypothesis that respondents neither agree nor disagree with this statement in favor of the alternative hypothesis that they agree with it ( $P < 0.01$ ,  $N = 93$ , Wilcoxon signed rank, two tailed).

At first pass, our findings appear to contradict those in previous field studies finding a strong positive effect of surveillance cues on altruistic behavior in natural environments [17-25]. Such interpretation of our findings however would be wrong. A critical difference between these studies and ours, stemming from the different research aims, is that the cues in these studies were placed in public spaces such as university cafeterias [17,18], public car parks [20], super markets [13,21] or hospital entrances [25], over an extended period of time. This implies that real reputational



concerns were at play. For example, participants in all these studies could self-select into several treatments, more than once, suggesting that individuals may be aware of the treatment manipulations and thus suspect they are being monitored. Similarly, since the manipulations occurred in places frequented by the participants, many of the encounters were likely to be neither anonymous nor one-shot, implying that reciprocity is not precluded by design. For these reasons, these studies suggest a potentially useful, low-cost, policy intervention (as was intended by the authors) but the evidence cannot inform the debate of whether altruistic behavior between strangers is maladaptive.

Taken together, the field evidence suggests that surveillance cues may be effective in promoting altruistic behavior in circumstances in which there are *real* opportunities to build a good reputation. In these instances, the cues may serve as a signal of what the expected behavior is and that behavior is monitored. In line with this is the finding that the surveillance-cue effect appears to be strongest when peer effects are modest [13,18], possibly due to the increased difficulty of monitoring behavior in large groups. Additional studies can help explore the underlying mechanism through which surveillance cues operate. Our findings indicate that surveillance-cues effects should not readily be interpreted as evidence that altruistic behavior between strangers is maladaptive.

## References

1. Fehr, E. & Fischbacher, U. (2003). The Nature of Human Altruism, *Nature* 425, 23 October 2003, 785-791.
2. Nowak, M.A. (2006). Five Rules for the Evolution of Cooperation. *Science* 314, 1560-1563.
3. Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology*, 64, 201–229.
4. Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology & Sociobiology*, 11, 375–424.
5. Haxby, J. V., Hoffman, E. A., & Gobbini, I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, 4, 223–233.
6. Izuma, K. (2012). The social neuroscience of reputation. *Neuroscience Research*, 72(4), 283–288.
7. Winston, J. S., Strange, B. A., O’Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, 5, 277–283.
8. Burnham, T. C., & Hare, B. (2007). Engineering human cooperation. *Human Nature*, 18(2), 88–108.
9. Burnham, T. C., & Johnson, D. D. P. (2005). The biological and evolutionary logic of human cooperation. *Analyse & Kritik*, 27(2), 113–135.
10. Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, 26, 245–256.

11. Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I.S., & Bateson, M. (2013). The watching eyes effect in the dictator game: It's not how much you give, it's being seen to give something. *Evolution and Human Behavior*, 34, 35–40.
12. Sparks, A., & Barclay, P (2013). Eye images increase generosity, but not for long: The limited effect of a false cue. *Evolution and Human Behavior*, 34, 317–322.
13. Ekström, M. (2012). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, 15, 530–546.
14. Fehr, E., & Schneider, F. (2010). Eyes are on us, but nobody cares: Are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences*, 277, 1315–1323.
15. Raihani, N.J., & Bshary, R. (2012). A positive effect of flowers rather than eye images in a large-scale, cross-cultural dictator game. *Proceedings of the Royal Society B: Biological Sciences*, 279, 3556–3564 .
16. Oda, R., Niwa, Y., Honma, A., & Hiraishi, K. (2011). An eye-like painting enhances the expectation of a good reputation. *Evolution and Human Behavior* 32, pp. 166–171.
17. Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2, 412–414.
18. Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior*, 32(3), 172–178.
19. Francey, D., & Bergmüller, R., (2012). Images of Eyes Enhance Investments in a Real-Life Public Good. *Plos One*. <http://dx.doi.org/10.1371/journal.pone.0037397>
20. Nettle, D., Nott, K., & Bateson, M. (2012) Cycle Thieves, We Are Watching You: Impact of a Simple Signage Intervention against Bicycle Theft. *Plos One*. <http://dx.doi.org/10.1371/>

[journal.pone.0051738](http://journal.pone.0051738).

21. Powell, K.L., Roberts, G., Nettle, D. (2012). Eye Images Increase Charitable Donations: Evidence From an Opportunistic Field Experiment in a Supermarket. *Ethology*. 118 (11), 1096-1101.
22. Bateson, M., Callow, L., Holmes, J.R., Roche M.L.R, & Nettle, D. (2013). Do Images of ‘Watching Eyes’ Induce Behaviour That Is More Pro-Social or More Normative? A Field Experiment on Littering. *Plos One*. <http://dx.doi.org/10.1371/journal.pone.0082055>.
23. Bateson, M., Robinson, R., Abayomi-Cole, T., Greenlees, J., O’Connor, A., Nettle, D., (2015) Watching eyes on potential litter can reduce littering: evidence from two field experiments. *PeerJ* 3:e1443 <https://doi.org/10.7717/peerj.1443>.
24. Fathi, M., Bateson, M., Nettle, D. (2014). Effects of watching eyes and norm cues on charitable giving in a surreptitious behavioral experiment. *Evolutionary Psychology*. 12 (5), 878-887.
25. King, D., Vlaev, I., Everett-Thomas, R., Fitzpatrick, M., Darzi, A., Birnbach, D., (2016). “Priming” hand hygiene compliance in clinical environments. *Health Psychology*, 35 (1), 96-101.
26. Pfoh, E., Dy, S., & Engineer, C. (2013). Interventions to improve hand hygiene compliance: brief update review. In: *Making Health Care Safer II: An Updated Critical Analysis of the Evidence for Patient Safety Practices*, Rockville. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK133371/>.
27. Pittet D, Allegranzi B, Sax H, et al. Evidence-based model for hand transmission during patient care and the role of improved practices. *Lancet Infectious Diseases*. 2006;6(10):641-652. doi:10.1016/S1473-3099(06)70600-4.

28. World Health Organization. (2009). Guidelines on Hand Hygiene in Health Care. Geneva: World Health Organization. Available at: <http://www.who.int/gpsc/5may/tools/9789241597906/en/>.
29. Cummings KL, Anderson DJ, Kaye KS. (2010). Hand Hygiene Noncompliance and the Cost of Hospital-Acquired Methicillin-Resistant Staphylococcus aureus Infection. *Infection Control & Hospital Epidemiology*. 31(4), 357-364.
30. Huis, A., van Achterberg, T., de Bruin, M., Grol, R., Schoonhoven, L., Hulscher, M. (2012). A systematic review of hand hygiene improvement strategies: a behavioural approach. *Implementation Science*. 7 (92). doi:10.1186/1748-5908-7-92.
31. Erasmus, V., Brouwer, W., van Beeck, E. F. et al (2009). A qualitative exploration of reasons for poor hand hygiene among hospital workers: Lack of positive role models and of convincing evidence that hand hygiene prevents cross-infection. *Infection Control and Hospital Epidemiology*, 30 (5), 415-419.
32. Borg, M. A., Benbachir, M., Cookson, B.D., et al (2009). Self-protection as a driver for hand hygiene among healthcare workers. *Infection Control and Hospital Epidemiology*, 30(6), 578-580.
33. Erasmus, V., Daha, T.J., Brug, H., et al. (2010). Systematic review of studies on compliance with hand hygiene guidelines in hospital care. *Infection Control and Hospital Epidemiology*. 31(3), 283-294. doi:10.1086/650451.
34. Northover, S., Pedersen, W.C., Cohen, A.B., & Andrews, P.W. (2017). Artificial surveillance cues do not increase generosity: two meta-analyses. *Evolution and Human Behavior* 38, 144–153.

35. Boyce J.M., & Pittet, D. (2002). Guideline for Hand Hygiene in Health-Care Settings. *Morbidity and Mortality Weekly Report*, 51, 1-45. Available at: <https://www.cdc.gov/mmwr/PDF/rr/rr5116.pdf>

### **Ethics statement**

All experimental protocols were approved by the Internal Review Board at New York University Abu Dhabi (#014-2016) and by the Social Science Research Ethics Committee at United Arab Emirates University (ERS-2015-3212). Informed consent was obtained by participants in a way which would not make participants aware that their behavior in the field experiment was studied (see section S1 in *SI*). Neither the existence nor the purpose of the field experiment that took place during the POSCE was revealed to participants at the end of the experiment. All aspects of our study, including the methods of consent and disclosure were carried out in accordance with relevant guidelines and regulations at United Arab Emirates University, and the IRB-approved protocols.

### **Data availability**

The data is submitted as a supplementary file for the revision process. If we are invited to revise our paper, we will submit our data to a permanent external data repository.

### **Competing financial interests**

The authors have no competing financial interests.

### **Author contributions**

N.N., E.K., A.D., and P.R. designed the experiment. E.K. and I.E. trained the research assistants. A.A. and I.E. oversaw the recruitment of participants. E.K., I.E., A.A., A.D., and N.N., supervised the experiment. E.K. compiled the dataset. A.D. and N.N. carried out the statistical analysis of the experimental data. N.N., A.D., and E.K. wrote the manuscript. N.N., E.K., A.D., and P.R. designed the survey. A.A. and I.E. administered the survey. E.K. coded the survey data. A.D. and N.N. analyzed survey responses.

## **Acknowledgements**

We thank the Editor, two anonymous reviewers, Loukas Balafoutas and Ernesto Reuben for useful comments on the paper, and Iain Blair for assisting us with the experiment. We also thank seminar participants at New York University in Abu Dhabi and at the Winter Experimental Social Sciences Institute at NYU Abu Dhabi, as well as the research assistants for helping with the experiment.



Supplementary Information for

**Surveillance cues do not enhance altruistic behavior  
among anonymous strangers in the field**

Erik J. Koornneef, Aurelie Dariel, Iffat Elbarazi,  
Ahmed R. Alsuwaidi, Paul B.M. Robben and Nikos Nikiforakis

This file includes the following sections:

- S1. Additional information on the experimental procedures and sample
- S2. Notes on the research assistants and robustness checks
- S3. Power calculations and sample size
- S4. The briefing note posted outside the room for the practice OSCE
- S5. The Observation Tool used by the RAs to evaluate the quality of hand hygiene
- S6. A picture of an 'OSCE room' used in the experiment
- S7. The pictures used in the experimental treatments
- S8. The WHO poster
- S9. Details of survey on hand hygiene
- S10. Survey and responses
- S11. Distribution of actions across treatments

## **S1. Additional information on the experimental procedures and sample**

The experiment was conducted in a large university ranking amongst the best in the Gulf Corporation Council region – the United Arab Emirates University – over two consecutive days in April 2016. The local population is of interest as for millennia locals interacted in small tribal groups with their personal reputation constantly at stake. This situation changed only in the last 20-30 years, when the discovery of vast oil reserves led to the development of large metropolises and a dramatically different lifestyle. If the evolutionary legacy hypothesis is correct, then we would expect the local population to be as sensitive to any surveillance cues as any other population studied previously.

The data collection occurred between 9 am and 3 PM on both days. A medical student who expressed interest in the POSCE was told to arrive at the venue at a specific time, such that interaction with other students was minimized. Once at the venue, students were asked to sign a consent form. The consent form informed individuals that, after the POSCE, they would be invited to respond to a survey on “Creating a better understanding of healthcare regulation: Exploring the experience, views and perceptions of postgraduate medical students in the United Arab Emirates”. While the consent form mentioned the POSCE, the focus throughout was placed on an anonymous survey that would be conducted afterwards on views on healthcare regulation. The POSCE was presented as an opportunity to build individual experience prior to the survey such that students were not concerned that they would be studied. After signing the form, the student was randomly assigned into one of four *private* examination rooms. Each practice slot and room was randomly assigned into one of three treatments. A total of 114 students of a population of 330 eligible students participated in the POSCE (response rate 34.2%). 90 participants were female (79%) and 24 were male (21%). This is similar to the gender composition of the student body at the medical program where 73.6% of all graduate students is female and 26.4% is male.

To minimize reputational concerns, the patient-examination rooms for the POSCE were located in an isolated wing of the university where students typically do not have access to. Each participating student was told to arrive on site at a specific time such that interaction with other students outside the examination rooms prior to the POSCE was minimized. Upon arrival, each student was randomly assigned into one of four *private* examination rooms. Each practice slot and room was in turn randomly assigned into one of three treatments.

## **S2. Notes on the research assistants and robustness checks**

The simulated patients were eight research assistants (RAs). The RAs were always of the same gender as the medical student, and were randomly reassigned to examination rooms over the course of the day.

All RA's completed a short training program, consisting of two online parts as well as a two-hour in person training. As part of the training program all participants successfully completed the online course from Hand Hygiene Australia (<http://www.hha.org.au/LearningPackage/olp-home.aspx>). The RA's also completed a special course prepared by the authors. As part of this course, they were asked to use an 'Observation Tool' to observe and evaluate the hand-hygiene quality of a minimum of 10 pre-recorded cases. On average, the RA's scored 95% of the observed behavior correctly. None of the RAs made contact with the participants before or after the experiment.

At any given point in the experiment, one RA waited for his/her blood pressure to be measured, and the other – seated at a faraway corner of the room – waited for his turn, while filling out a Sudoku book. In actuality, this individual was covertly monitoring the student's HH practice. This implies that only one RA was responsible for recording a given observation. While multiple coders may have been desirable we decided against it as this would require adding more RAs in the room which could make students more self-aware. Importantly, however, the existence of one coder per observation should not bias our findings as there were several RAs overall, and each recorded roughly the same number of observations for all treatments at different points in time. Therefore, any systematic biases on the RA level (e.g., one being stricter in evaluating than another) should have no effect on our treatment estimates.

With one exception, all research assistants were unknown to the students and instructed not to share any personal information with participants. Due to being short of one RA, at the last minute, we recruited as an RA a staff member who may have been familiar to some of the participants. However, as one would expect given that RAs were randomly assigned to treatments, our results are unaffected if we exclude observations from this RA. In particular, with regards to the time spent washing, when using observations from all RAs, the average time spent washing was 21.5 seconds, and we find no statistically significant differences across treatments (Eyes vs. Baseline:  $P=0.69$ ,  $N=71$ ; Camera vs. Baseline:  $P=0.64$ ,  $N=79$ ; Mann-Whitney Test, two-tailed). Excluding the one RA who may have been known to some participants, the average time spent washing was 21 seconds, and again we find no statistically significant differences across treatments (Eyes vs. Baseline:  $P=0.78$ ,  $N=60$ ; Camera vs. Baseline:  $P=0.90$ ,  $N=64$ ; Mann-Whitney Test, two-tailed). With regards to the quality of hand coverage, when using observations from all RAs, there were

no significant differences across treatments (Eyes vs. Baseline:  $P=0.99$ ,  $N=71$ ; Camera vs. Baseline:  $P=0.18$ ,  $N=79$ ; Mann-Whitney Test, two-tailed). Excluding the one RA who may have been known to some participants, we again find no significant treatment differences (Eyes vs. Baseline:  $P=0.76$ ,  $N=60$ ; Camera vs. Baseline:  $P=0.66$ ,  $N=64$ ; Mann-Whitney Test, two-tailed). Finally, with regards to the extent to which participants across treatments turned off the tap after washing their hands using a paper towel, we find no treatment differences whether we use observations from all RAs, (Eyes vs. Baseline:  $P=0.39$ ,  $N=71$ ; Camera vs. Baseline:  $P=0.90$ ,  $N=79$ ; Mann-Whitney Test, two-tailed) or if we exclude observations from that one RA (Eyes vs. Baseline:  $P=0.12$ ,  $N=60$ ; Camera vs. Baseline:  $P=0.77$ ,  $N=64$ ; Mann-Whitney Test, two-tailed).

The main role of the RAs for our research purposes was to accurately record the hand hygiene quality of the medical students prior to treating the patient. Once the student was in front of the washbasin, the RAs recorded the duration of the hand washing as well as the level of compliance with the recommended technique, using the Observation Tool. Although RAs were instructed to appear as bored and indifferent as possible, they were asked to gently prompt participants to wash their hands prior to taking their blood pressure in case the participants didn't first wash their hands. RAs were instructed to do so in the most casual manner (e.g., by saying "Would you mind washing your hands first?"). This was essential for participants to be exposed to the experimental treatment, and occurred in 54 of the 114 instances (47%). As this could in principle activate participants' reciprocity-based psychology, we repeat here the analysis using only observations from unprompted participants.

The likelihood a participant had to be prompted, as one might have expected, is unrelated to the treatment manipulation ( $P=0.77$ ,  $\chi^2=0.68$ ). Accordingly, our results are unaffected if we only use observations in which participants washed their hands without being prompted. In particular, we find no statistically significant differences across treatments when comparing the time spent washing (Eyes vs. Baseline:  $P=0.91$ ,  $N=36$ ; Camera vs. Baseline:  $P=0.49$ ,  $N=36$ ; Mann-Whitney Test, two-tailed), the quality of hand coverage (Eyes vs. Baseline:  $P=0.64$ ,  $N=36$ ; Camera vs. Baseline:  $P=1.00$ ,  $N=36$ ; Mann-Whitney Test, two-tailed), and the extent to which participants across treatments turned off the tap after washing their hands using a paper towel (Eyes vs. Baseline:  $P=0.14$ ,  $N=36$ ; Camera vs. Baseline:  $P=0.88$ ,  $N=36$ ; Mann-Whitney Test, two-tailed).

### **S3. Power calculations and sample size**

Convention prescribes that the sample size is such that a given treatment effect will be significant at the 5% level, 80% of the time (that is, the power of the test is 80%). Our study is the first to explore the impact of surveillance cues on the quality of hand hygiene in a natural field setting. We therefore have no prior on the potential size of the treatment effect (if any). To ensure that our statistical tests would be sufficiently powered we used the following approach to calculate our sample size. As a benchmark for the average quality of hand hygiene in our baseline condition we used the average compliance rate across a number of observational studies, which is 40%<sup>35</sup>. As mentioned, the photograph of the male eyes in our experiment was taken from (23) who found an effect of 122%. Given these numbers, we need a sample of 15 participants per treatment to detect a significant difference at the 5% level, 80% of the time. Given our sample, we have a power of 99.72%.

#### **S4. Briefing note posted outside room**

### **Practice OSCE**

#### ***Blood Pressure Measurement***

#### **AIM**

The aim of this station is to assess your ability to measure blood pressure correctly whilst communicating appropriately with the patient.

#### **TASK**

Measure the patient's blood pressure. There is no need to take a history from the patient, but you may ask questions and give any explanations and advice that you feel are appropriate.

#### **SCENARIO**

You do not know this patient and there is nothing of significance in his/her records.

#### **PATIENT**

There will be two simulated patients in the room. They will take it in turns to be the "patient".

#### **EQUIPMENT**

- An aneroid sphygmomanometer and a mercury sphygmomanometer.
- A standard cuff (size 23cm x 12 cm)
- A stethoscope

#### **WHEN YOU HAVE FINISHED**

You may tell the patient what the blood pressure is.

**S5. The Observation Tool used by the RAs to evaluate the quality of hand hygiene**

<b>Clinical Skills Observation Tool - Blood Pressure Measurement</b>				
		Participant ID:		
		Research Assistant ID:		
<b>Part 1: Preparation</b>		<b>Performed adequately &amp; completely</b>	<b>Attempted but inadequate / incomplete</b>	<b>Not attempted</b>
1	Introduces self (no identifying information must be given)			
2	Checks that the patient has been resting for 3 minutes and has not eaten for 30 minutes, if yes, ask patient for their consent to carry out the procedure			
<b>Part 2: Hand hygiene</b>		<b>Performed adequately &amp; completely</b>	<b>Attempted but inadequate / incomplete</b>	<b>Not attempted</b>
3	Participant needed a prompt/reminder to wash hands.	YES / NO		
4	Coverage – are all hand surfaces cleaned adequately (i.e. did the participant adequately follow all the steps)			
5	Dries hands with paper towel.			
6	Turns off tap with paper towel			
7	Throws paper towel in correct bin	YES / NO		
8	<b>Total time spent washing hands (in seconds, round up to the nearest 5 sec.)</b> <small>Note: start clock when tap is turned on and stop clock when tap is turned off</small>			
<b>Part 3: Blood pressure measurement</b>		<b>Performed adequately &amp; completely</b>	<b>Attempted but inadequate / incomplete</b>	<b>Not attempted</b>
9	Assists the patient into a comfortable position. Ensure that the arm is at heart level and resting on a suitable firm surface			
<b>Part 4: Hand hygiene</b>		<b>Performed adequately &amp; completely</b>	<b>Attempted but inadequate / incomplete</b>	<b>Not attempted</b>
10	Participant needed a prompt/reminder to wash hands.	YES / NO		
11	Coverage – are all hand surfaces cleaned adequately (i.e. did the participant adequately follow all the steps)			
12	Dries hands with paper towel.			
13	Turns off tap with paper towel			
14	Throws paper towel in correct bin	YES / NO		
15	<b>Total time spent washing hands (in seconds, round up to the nearest 5 sec.)</b> <small>Note: start clock when tap is turned on and stop clock when tap is turned off</small>			
<b>Additional Comments/Observations:</b>				
Please circle one number to indicate your level of agreement with the following statement regarding the candidate’s performance:				
“The candidate demonstrated a professional approach that included concern for the patient and maintained patient dignity”				
<b>Strongly agree</b>		<b>Strongly disagree</b>		
5	4	3	2	1




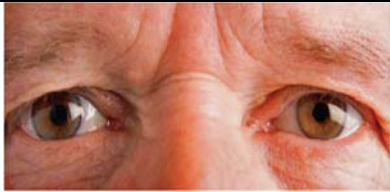

**S6. A picture of an ‘OSCE ROOM’ used in the experiment**



**S7. The pictures used in the experimental treatments**

---

---

Baseline	Eyes	Camera
		

---

---

## S8. The WHO poster

# How to Handwash?

WASH HANDS WHEN VISIBLY SOILED! OTHERWISE, USE HANDRUB

 Duration of the handwash (steps 2-7): 15-20 seconds

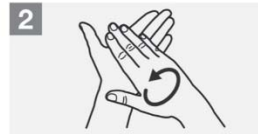
 Duration of the entire procedure: 40-60 seconds



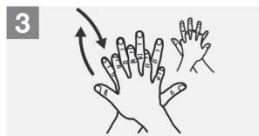
Wet hands with water;



Apply enough soap to cover all hand surfaces;



Rub hands palm to palm;



Right palm over left dorsum with interlaced fingers and vice versa;



Palm to palm with fingers interlaced;



Backs of fingers to opposing palms with fingers interlocked;



Rotational rubbing of left thumb clasped in right palm and vice versa;



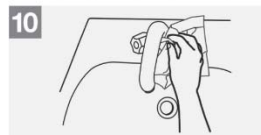
Rotational rubbing, backwards and forwards with clasped fingers of right hand in left palm and vice versa;



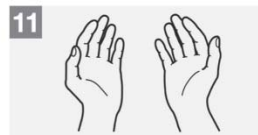
Rinse hands with water;



Dry hands thoroughly with a single use towel;



Use towel to turn off faucet;



Your hands are now safe.



World Health Organization

Patient Safety

A World Alliance for Safer Health Care

SAVE LIVES

Clean Your Hands

Based on the 'How to Handwash', URL: [http://www.who.int/gpsc/5may/How\\_To\\_HandWash\\_Poster.pdf](http://www.who.int/gpsc/5may/How_To_HandWash_Poster.pdf) © World Health Organization 2009. All rights reserved

May 2009

## **S9. Details of the survey on hand hygiene**

To investigate whether hand hygiene prior to treating a patient in our experiment is an altruistic act we administered a survey to 100 medical students with the same level of training as those who participated in our experiment. Respondents were presented with a vignette designed to mimic the situation and incentives in our experiment, and asked whether they would wash their hands prior to treating the patient or not, and the reasons for their decisions. The survey also included a question to evaluate our claim that the quality of hand-hygiene matters when it comes to reducing infection by asking participants whether they agree that washing hands for longer reduces the risk of infection for the patient. The full survey is available in S10.

The survey was administered in the last week of February 2018. Respondents were students in preclinical years, from years 3-5 of the MD Program. They were approached either in their classes (7 classes) or in the clinical lab skills sessions (4 sessions) and were invited to fill in the questionnaire. Year 5 students were approached in their clerkship rotation. Those who agreed to participate were provided with the survey and instructed to drop the filled-out survey at a closed box outside the office of the education coordinator. Most students agreed to participate. The few who chose not to participate had no information about the content of the survey.

The survey presented students with a simple vignette designed to mimic the situation and incentives in our experiment. To avoid confusing respondents who were still unfamiliar with the OSCE, we refrained from referring to either an OSCE or a POSCE in the vignette, placing instead the interaction in a medical environment that would be familiar to all medical students. Respondents were then asked whether they would wash their hands prior to treating the patient or not, and the reasons for their decisions. The different responses appeared in a random order.

The responses indicate clearly that the main concern driving hand hygiene prior to treating a patient appears to be a desire to do no harm to him/her. We can reject the hypothesis that respondents neither agree nor disagree with the statement that they would wash their hands to avoid harming the patient in favor of the alternative hypothesis that they agree with this reason (Wilcoxon,  $p < 0.01$ ,  $N=87$ ). Participants are significantly more likely to agree with this reason than with any of the other reasons (Wilcoxon,  $p < 0.01$ ,  $N=87$ , for all pairwise comparisons). Only 12.4% of respondents who said they would wash their hands before taking the patient's blood pressure had a reason with which they agreed more strongly than avoiding harm.

The survey also included a question to evaluate our claim that washing hands for longer is believed to reduce the risk of infection. We find clear evidence that respondents believe the duration of HH is an important determinant of the likelihood of infection. We can reject the hypothesis that

respondents neither agree nor disagree with the statement that “A doctor who washes her hands for 30 seconds is less likely to infect a patient than a doctor who washes her hands for 10 seconds” in favor of the alternative that they agree (Wilcoxon,  $p < 0.01$ ,  $N = 87$ ).

## S10. Survey and responses

### Welcome

We are a team of researchers from United Arab Emirates University, New York University in Abu Dhabi and Erasmus University in Rotterdam. We are conducting a short survey regarding hand hygiene among practitioners. Please spare us a few minutes of your time to answer 3 short questions.

### Consider the following situation

You have a 10-minute appointment with a patient in your office. This patient is not one of your regular patients. The patient normally sees another doctor, but today this doctor is unavailable. There is nothing of significance in the patient's record. Since this is the first time you see the patient you decide to take her/his blood pressure. You have just arrived at your practice for the appointment after driving thirty minutes from your home.

### Question 1 (N=100)

Would you wash your hands **before** taking the patient's blood pressure? (tick one answer)

Yes 87%

No 13%

If you answered "Yes", please go to page 2, otherwise, please go straight to page 3.

**Question 2 [Only for those who answered YES]: (N=87)**

Please state how much you agree with the following reasons for washing your hands **before** taking the patient's blood pressure by ticking one of the answers.

[Note: entries are the percentage of respondents who gave a particular answer, except in the last column which presents the mean response. \*\*\* and \*\* indicate that a Wilcoxon test rejects the null hypothesis that the mean is not different from 0 at the 1% and 5% level, respectively]

	Strongly disagree (-2)	Somewhat disagree (-1)	Neither agree nor disagree (0)	Somewhat agree (1)	Strongly agree (2)	Mean response
1. I wash my hands because I wish to do no harm to the patient by infecting them	0.00%	1.15%	2.30%	12.64%	83.91%	1.79***
2. I wash my hands because that is what all other clinical practitioners do	13.95%	15.12%	39.53%	26.74%	4.65%	-0.07
3. I wash my hands because it reduces the likelihood the patient complains about me	25.58%	15.12%	29.07%	24.42%	5.81%	-0.30**
4. I wash my hands because that is what all clinical practitioners are expected to do	5.81%	3.49%	19.77%	36.05%	34.88%	0.91***
5. I wash my hands because I want to be a good ambassador for my practice/hospital	2.33%	2.33%	9.30%	24.42%	61.63%	1.41***

Please state any other reasons for washing your hands. \_\_\_\_\_  
 (Note: 26.4% of respondents provided responses here, however, the vast majority of them simply provided variants of the “do no harm” response, suggesting they used the opportunity to elaborate on their reasons for washing their hands.)

**Question 2 [Only for those who answered NO]: (N=13)**

Please state how much you agree with the following reasons for **not** washing your hands **before** taking the patient's blood pressure by ticking one of the answers.

*[Note: entries are the percentage of respondents who gave a particular answer, except in the last column which presents the mean response. \*\*\* and \*\* indicate that a Wilcoxon test rejects the null hypothesis that the mean is not different from 0 at the 1% and 5% level, respectively]*

	Strongly disagree (-2)	Somewhat disagree (-1)	Neither agree nor disagree (0)	Somewhat agree (1)	Strongly agree (2)	Mean response
1. Not washing my hands will not harm the patient	23.08%	15.38%	30.77%	15.38%	15.38%	-0.15
2. Not washing my hands will reduce the consultation time	53.85%	15.38%	7.69%	23.08%	0.00%	-1**
3. Not washing my hands speeds up the process of building immunity to disease for me	30.77%	15.38%	46.15%	0.00%	7.69%	-0.62*
4. Washing my hands irritates my skin/	53.85%	15.38%	15.38%	15.38%	0.00%	-1.08**

Please state any other reasons for not washing your hands. \_\_\_\_\_

(Note: 76.9% of respondents provided responses here. The majority of them said that hand washing was not necessary and hand sanitization was sufficient. Among the others, some mentioned that hand washing was not necessary, time pressure or that they would forget.)



**Question 3 [for all respondents]: (N=93)**

Please state how much you agree with the following statement by ticking one of the answers:

*“A doctor who washes her hands for 30 seconds is less likely to infect a patient than a doctor who washes her hands for 10 seconds.”*

---

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	Mean
3.23%	7.53%	10.75%	39.78%	37.63	1.01***

---

*[Note: entries are the percentage of respondents who gave a particular answer, except in the last column which presents the mean response. \*\*\* and \*\* indicate that a Wilcoxon test rejects the null hypothesis that the mean is not different from 0 at the 1% and 5% level, respectively]*

**Thank you!**

**S11. Distribution of actions across treatments**

