Measuring the moral impact of operating "drones" on pilots in combat, disaster management and surveillance

Christen, Markus ; Michael, Villano ; Darcia, Narvaez ; Jesùs, Serrano

Abstract: Remotely piloted aircrafts (RPAs or "drones") have become important tools in military surveillance and combat, border protection, police and disaster management. In particular, the use of weaponized RPAs has led to a discussion on the ethical, strategic and legal implications of using such systems in warfare. In this context, studies suggest that RPA pilots experience similar exposure to post-traumatic stress, depression and anxiety disorders compared to fighter pilots, although the flight and combat experiences are completely different. In order to investigate this phenomenon, we created an experiment that intends to measure the "moral stress" RPA pilots may experience when the operation of such systems leads to human casualties. "Moral stress" refers to the possibility that deciding upon moral dilemmas may not only cause physiological stress, but may also lead to (unconscious) changes in the evaluation of values and reasons that are relevant to problem solving. The experiment includes an RPA simulation based on a game engine and novel measurement tools to assess moral reasoning. In this contribution, we outline the design of the experiment and the results of pretests that demonstrate the sensitivity of our measures. We close by arguing for the need of such studies to better understand novel forms of human-computer interaction.

Originally published at:
Christen, Markus; Michael, Villano; Darcia, Narvaez; Jesùs, Serrano (2014). Measuring the moral impact of operating "drones" on pilots in combat, disaster management and surveillance. In: 22. European Conference on Information Systems, Tel Aviv, 9 June 2014 - 11 June 2014, 1-10.

MEASURING THE MORAL IMPACT OF OPERATING "DRONES" ON PILOTS IN COMBAT, DISASTER MANAGEMENT AND SURVEILLANCE

Research in Progress

Christen, Markus, University of Zurich, Zurich, Switzerland, christen@ethik.uzh.ch

Villano, Michael, University of Notre Dame, Notre Dame IN, USA, mvillan1@nd.edu

Narvaez, Darcia, University of Notre Dame, Notre Dame IN, USA, dnarvaez@nd.edu

Jesùs Serrano, University of Notre Dame, Notre Dame IN, USA, jgs121@gmail.com

Charles R. Crowell, University of Notre Dame, Notre Dame IN, USA, ccrowell@nd.edu

Abstract

Remotely piloted aircrafts (RPAs or "drones") have become important tools in military surveillance and combat, border protection, police and disaster management. In particular, the use of weaponized RPAs has led to a discussion on the ethical, strategic and legal implications of using such systems in warfare. In this context, studies suggest that RPA pilots experience similar exposure to post-traumatic stress, depression and anxiety disorders compared to fighter pilots, although the flight and combat experiences are completely different. In order to investigate this phenomenon, we created an experiment that intends to measure the "moral stress" RPA pilots may experience when the operation of such systems leads to human casualties. "Moral stress" refers to the possibility that deciding upon moral dilemmas may not only cause physiological stress, but may also lead to (unconscious) changes in the evaluation of values and reasons that are relevant to problem solving. The experiment includes an RPA simulation based on a game engine and novel measurement tools to assess moral reasoning. In this contribution, we outline the design of the experiment and the results of pretests that demonstrate the sensitivity of our measures. We close by arguing for the need of such studies to better understand novel forms of human-computer interaction.

Keywords: *Remotely piloted aircraft, moral decision making, combat, human-computer-interaction in stress situations.*

1 Introduction

1.1 The "Rise of the Drones"

Since the beginning of human aviation, the idea of airplanes without human pilots has been considered. But it took several decades until unmanned flight became a relevant factor in aviation beyond recreational use of model aeroplanes. Remotely Piloted Aircrafts (RPAs) – the preferred name given by the U.S. Air Force to clarify that a human pilot controls the device – or Unmanned Aerial Vehicles (UAVs) have entered the public discourse with the label of "drones." These systems are increasingly used in various applications – mainly for surveillance, but also as instruments to execute deadly force in the war on terror following the 9/11 terrorist attack, particularly emphasized by the Obama administration (McCrisken, 2013). A major line of supportive moral argumentation of “drone strikes” is that they do not endanger the life of a pilot and are more precise and thus diminish the moral dilemmas associated with aerial warfare, e.g. related to strategic bombing (Schaffer 1988).

Most of the current discussion regarding “drones” focuses on weaponized RPAs and their ethical, legal and strategic implications (Enemark, 2013; McCrisken, 2013; Sauer and Schörnig, 2012). Although only three countries are known to have used RPAs in combat to date (USA, Israel, United Kingdom), more than 75 may possess some type of RPA (Kmietowicz, 2012). This shift towards “drones” seems to be an inevitable evolution in a time where asymmetric war against hidden enemies dominates warfare and military cost restrictions dictate budget planning. From 2007 to 2010, the budget requests for RPAs of the U.S. Department of Defense (USDD) increased by more than 50% (Shaw and Akhter, 2012) and the USDD plans to spend $30.8 billion more on developing and acquiring RPAs between 2011 and 2015 (Miller, 2012). One example is the so-called “Predator” that costs $20 million for the entire system (consisting of four aircraft, a ground station, a satellite link and a maintenance crew) and is used for armed reconnaissance and surveillance. It has a wingspan of 14.8 meters (48.7 feet), a length of 8.2 meters (27 feet), and weighs 512 kg (1130 pounds). It carries two laser-guided AGM-114 Hellfire missiles, can fly up to 7620 meters (25’000 feet), and is piloted by three personnel: a pilot, a sensor operator, and a mission intelligence coordinator (USAF, 2009). Its successor, the “Reaper” RPA has a larger payload capacity and outperforms the “Predator” also regarding the other capabilities.

However, a focus on military should not overlook the various non-military applications of RPAs; potential markets range from law enforcement to wildlife management and oil discovery. For example, an unmanned Predator B drone, originally designed to gather intelligence for the military, provided firefighters with up-to-the-minute data on the many conflagrations in a large Californian forest fire (Cohen, 2007). Drone technology may become widespread also in the civilian market. The market research organization Teal Group estimates that RPA spending will almost double over the next ten years from current worldwide sales of US$5.9 billion annually to US$11.3 billion (Schneiderman, 2012).

1.2 Human-RPA-Interaction

The development of RPAs has led to an expanded number of available platforms in the unmanned vehicle sector: micro air vehicles, mini UAVs, tactical RPAs, medium-altitude/long-endurance RPAs, and high-altitude/long-endurance RPAs. A significant aspect is their very high electronic content, which requires a broad spectrum of skills. Therefore, it has become clear that RPAs create a demand for a new type of pilot – a combination of aeronautical engineer, physicist, and software coder. Optimizing human control of RPAs has become a significant topic in human factor research. One topic is automation of some aspects of navigation, where it has been shown that reliable automation can help alleviate task interference and reduce workload, thereby allowing pilots to better handle concurrent tasks during single- and multiple-RPA flight control (Dixon et al., 2005). Another topic is operator capacity (Cummings and Guerlain, 2007), improvement in signal detection in vigilance (Gunn et al., 2005) and sonification (continuous auditory alerts mapped to the state of the monitored task) to decrease the “physical detachment” of the pilot from the mission (Donmez et al., 2009). This brief overview demonstrates that most research on human-RPA interaction concerns the efficiency of RPA operation and thus capacities like executive function and cognition.

1.3 Conceptualizing “moral stress” of RPA pilots

Recent findings suggest that the focus on cognitive and executive functions may not be sufficient in order to completely understand the human-RPA interaction. Several studies indicate an unforeseen impact on RPA pilots. For example, the US Air Force conducted a survey that showed that 46% of Reaper and Predator pilots and 48% of Global Hawk sensor operators (surveillance RPA) suffered from ‘high operational stress’ (vaguely defined in the study). A smaller but significant number of operators had ‘clinical distress’: anxiety, depression or stress severe enough to affect an operator’s job performance or family life (Bumiller, 2011). RPA operators also experience physical exhaustion, and have been referred to as the most fatigued flight crews in the military (Trimble, 2008). There are reports on RPA crews suffering from post-traumatic stress induced by constant exposure to high-resolution images of real-time killing and the after-action inventory of body parts (Lindlaw, 2008). A study by the US Armed Forces Health Surveillance found that, among RPA pilots, the incidence of stress disorders is similar to those who pilot manned aircraft (Otto, 2013; Ortega, 2013).

So far, there is only speculation about the causes of specific psychological problems of RPA pilots. A major line of argumentation focuses on the different circumstances of RPA pilots in comparison to traditional combat roles. In particular, RPA pilots have to deal with a lack of deployment rhythm and of combat compartmentalization (i.e., a clear demarcation between combat and personal/family life); fatigue and sleep disturbances related to shift work; austere geographic locations of military installations supporting RPA missions; social isolation during work, which can diminish unit cohesion and thereby increase susceptibility to PTSD; and sedentary behavior with prolonged screen time (Otto, 2013). Thus, there is a difference between what military RPA operators actually do and what they – and society – perceive soldiers do (Medact, 2013). Previous soldiers had to overcome risk, fear, pain and discomfort. They were held in respect especially if the war was perceived as ‘just’. The fact that their personal needs were subjugated to the common good also enhanced their self-respect. But RPA operators have been described as ‘outsourcing risk’ (Singer 2009). The aspects of battle that normally enhance self-esteem and engender the esteem of others are absent so there is the potential for the work to erode self-image as well as social image as “war heroes” in the public mind. RPA pilots also lack the camaraderie of fellow soldiers and the opportunity to share feelings based on common experience (Singer, 2009).

These factors partly may explain the findings that RPA pilots experience more psychological disorders than otherwise would be expected. However, we suggest that an additional mechanism is at work, which we conceptualize as “moral stress,” which refers to the moral content of the actions performed by RPA pilots. In other words, the interaction between pilot and the actual mission goals, rather than between pilot and RPA, is relevant. This view contradicts the claimed similarity between flying drones and playing video games. Air Force officials, for example, make statements like: “when you’re looking at a skill set for what it takes to be able to work in a two-dimensional area, gaming helped me to make the transition” (Lieutenant Colonel Geoff Barnesquoted, quoted in Mockenhaupt, 2009). But we claim that this analogy is misleading with respect to the moral impact of the actions on the pilots. Along with Gregory (2011), we believe that RPA operations do not reduce war to a video game in which the killing space appears remote and distant, but rather that this new war strategy with the target’s constant visibility produce a special kind of intimacy. There already has been some speculation about the effect on drone operators of observing the target they have been assigned to kill on a high definition computer screen. Tracking targets for long periods of time may enhance empathy with them, rather than create a sense of detachment. But exactly how this process impacts on the operator’s perception of the target as a human being is unknown (Lee, 2012).

Our study intends to inform this debate by analyzing changes in moral reasoning following necessarily tragic decisions an RPA pilot has to make albeit in a simulated environment. We conceptualize “moral stress” as being involved in decisions with high moral relevance (killing of people) without physiologically experiencing the situational factors that allow for “dealing” with the consequences. Thus, “moral stress” is not equivalent to “moral distress”, which is the psychological disequilibrium and negative feeling state experienced when a person makes a moral decision but does not follow through by performing the moral behavior indicated by that decision (Wilkinson 1987/88). Rather, it concerns the experience of a dilemma in which any decision violates important moral values. This understanding of "moral stress" has been developed in the context of nursing ethics (Lützén et al. 2003), where dilemmas between competing values (e.g., autonomy vs. harm prevention) are quite common in health care. How exactly stress is measured in such situations is controversial, as stress generally involves physiological, psychological and social components (Monat & Lazarus 1985). In addition, it is generally known that a threatening situation that instigates a stress response can change what types of rationale and reasoning seem applicable (Pyszczynski et al., 2003). Therefore, our notion of "moral stress" does not only concern the fact that decisions of high moral relevance cause physiological, psychological or social stress, but that these decisions have the (unconscious) potential to change the evaluation of values and reasons that are relevant for the decision problem. Our methodology thus involves not only measures for known stress components, but also includes novel measurement tools to assess moral reasoning.

In our experiment, we certainly cannot replicate the real decisions that have deadly consequences for the victims of drone strikes. We therefore have chosen a well-studied tragic dilemma – the Trolley Dilemma, where people can benefit one or more persons at the cost of harming others (Foot, 1967; Thomson, 1985) – and we have embedded it in three scenarios (military, disaster management, surveillance) where RPA pilots operate a simulated drone in defined missions. We hypothesize that tragic decisions that are embedded in a military context (compared to other contexts) will change moral reasoning (captured by value and reason evaluations) more strongly and will be more strongly associated with stress measures both during the operations as well as justifications of actions during a mission debriefing. Below, we outline the experimental design and present results of pre-tests indicating that the measures we employ should be sensitive with respect to the study goals.

2 Methods

2.1 Experimental Design

Our study consists of a between-subject-design that includes five steps outlined in Table 1. Participants will be members of the Reserve Officer Training Corps (branches: Air Force, Army, Navy). The goal was to recruit participants who may have a comparable mind-set and age as military RPA pilots, since military RPA pilots are increasingly recruited directly from university programs (e.g. the University of North Dakota was the first to offer a specific degree program in 2009; Dillow 2013) and are likely to have pro-military attitudes. Certainly, experienced RPA pilots may still show different reactions when confronted with our simulated dilemma. However, we are currently not in the position to recruit a sufficient number of military RPA pilots for study purposes.

Our participants will have the role of the RPA *pilot*, i.e. they will guide the RPA to the optimal launch point of the missile and they will keep the target in crosshairs. They do not decide whether the target is legitimate (task of the sensor operator) nor do they launch the simulated weapon, a missile (task of the mission intelligence coordinator); however, they can shift the missile target to an alternative target (resulting in collateral damage), which is implemented in our experiment by pressing a button. After impact, the pilots will have to approach the scene and report the damage made, before proceeding to the next mission.

The whole experiment takes part in an acoustic booth that includes a large 60' LED monitor to ensure immersion. Participants will wear a headset to obtain automatized orders and to respond to orders using the microphone (answers will be recorded). They will use a joystick to operate the RPA, and a computer mouse to complete questionnaires. For measuring the physiological stress response, a finger pulse-ox monitor will be attached to the index finger of the non-dominant hand, which will be slightly restrained to minimize movement artifacts. Heart rate data will be examined through analysis of the photoplethysmograph (PPG) waveform, a graphical representation of the measured change in blood flow volume during the cardiac cycle (Shelley, 2007). This measurement is most sensitive in the fingers and detects changes in the sympathetic nervous system (Awad et al., 2001). Pre-stimulus baseline heart rate will be determined by the heart rate during the last second prior to each targeted event. For each 1-second interval, the mean heart rate will be calculated. Using procedures recommended in Gamer et al. (2006/2008), post-stimulus heart rate difference scores will be calculated by subtracting the pre-stimulus baseline heart rate from each of the 20 1-second post-event heart rates. From these results, maximum heart rate deceleration, if any, will be calculated as well as an average of all heart rate differences from baseline.

Step 1	Preparatory phase: informed consent, briefing, installation of the participant in the experimental box		
Step 2	Survey part 1: General information, state and trait measures		
Step 3	Scenario 1: Military combat	Scenario 2: Disaster management	Scenario 3: Surveillance
Step 4	Survey part 2: reason and value test, state measures		
Step 5	Debriefing using a semi-structured interview		

Table 1. Outlining the five steps of the experimental procedure: The physiological stress response is continuously measured in steps 2-5.

Participants will be randomly assigned to one of the three groups each experiencing a different scenario. We believe that these three scenarios will influence moral reasoning differently. The simulation part lasts ~30-40 minutes and includes the following steps: First, an introduction video explains the general framework and intends to prime the participants for either scenario (military, disaster management, surveillance). Second, in the flight training, the participant will start the RPA and follow the commands to designated landmarks. Third, in mission training, the participant will start the RPA and then be directed to the training place, where training suitable to the missions (see section 2.2.) will take place. For example, in the military setting, the participant will first hit the target; second they will be ordered to redirect the missile to an alternative target; and third, they will be given the choice to redirect the missile in a short amount of time (10 seconds). Finally, the participant will be guided to five missions. Mission 2, 4 and 5 are dilemmatic, i.e. require to choose between doing nothing (five casualties) and redirecting the missile (one casualty, but the "value" of the single victim increases in each mission from being the same as the other five persons, to being a commander, to being a family member).

Based on current knowledge (Bleske-Rechek et al., 2010), it is expected that the participant will choose to redirect in mission 2 but not in mission 5. The actual decisions, however, are not the primary interest of the study, but rather of central interest will be any changes in reasoning and value assessment due to the experience of the simulation.

In pre-tests, we have obtained baseline data for the test battery and we checked the usability of the whole experimental setting. In this research-in-progress paper, we only report results of these pre-tests.

2.2 Simulation environment

The simulation was created using Valve Software's Source SDK Base 2013 game engine (Valve, 2013) and development tools. A Half-Life 2 MOD (game modification) was developed using the Hammer editor to construct the 3D world, build custom models, and develop much of the gaming logic. Additional gaming logic was developed directly in the available Source SDK C++ code freely available through GitHub. Source provided a robust development environment which was adapted to produce a drone simulation by creating an invisible floor above the landscape on which the "player" walks which creates the illusion of flight from a RPA. Take off and landings are simulated by the use of an invisible ramp. Demonstrations of the simulator to experienced video game players have found the effect of flying a drone to be very compelling. The participant controls the forward movement and "player look" through a joystick, and can redirect a missile in the moral dilemma scenarios by pressing a joystick button (button 2 on a Logitech Attack 3 joystick). To simulate a switch to a targeting camera, (and thus control the participant's view of the missile attack to insure they see what the experimenters want them to see) the "point_viewcontrol" entity was programmed in the game map and tracked along a path to slowly zoom in on the target area. Once an attack is completed, the view and movement control is returned to the participant. All commands and prompts are given by an automatized system using both visual (text output on the screen) and audio modalities (prepared audio files, triggered during game play). Prompts are integrated into the simulation in order to avoid large deviations from mission paths, whereas commands are triggered when the RPA has reached a designated spot.

The simulation includes three different scenarios, each experienced by a different group of subjects:

a) Combat scenario: The narration framing the scenario is a country involved in a war against terror. The RPA pilot mission is to find groups of terrorists following commands given by the simulation. When the drone reaches a defined spot, a missile is fired towards a group of five persons (suspected terrorists). The only decision the pilot can make is to redirect the missile to another target within a short time span, if intelligence indicates that the original target group has been misidentified as terrorists. Redirecting will lead to a collateral damage in the alternative target area (one person will be killed).

b) Disaster management scenario: The narration framing the scenario is a country threatened by large forest fires. The RPA pilot mission is to find so-called high intensity fires that have to be hit by water bombs mounted on missiles to extinguish the fire (fictitious adaptation of the existing 2RS system, an explosive hose firefighting system; CORDIS 2004). Again, the pilot will follow commands given by the simulation to direct the RPA to the launch point of the missiles. In some missions, five firemen will suddenly appear on scene and would be hit by the missile, if no decision is made to redirect it. As in the military scenario, redirecting the missile involves collateral damage.

c) Surveillance scenario: The narration framing the scenario is conducting a census in a national park. The RPA pilot mission is to find groups of persons in the park and photograph them. The only decision required from the pilot in 3 out of five missions is to choose which groups should be photographed. This scenario is essentially a control scenario.

The scenarios should allow disentangling the effect of immersion alone (surveillance) from different intentional settings: in the military setting, the pilot is involved right from the beginning in a "kill mission", whereas in the disaster management setting, the tragic decision results from a pure accident. Although we hypothesize that the immediate physiological stress response across participant groups will be the same in both scenarios, we suspect that the stress responses will differ when the actions will have to be justified in the debriefing, and we also expect that the reasoning and value evaluations will differ due to the higher probability that military-related tragic decisions lead to psychological damage; a phenomenon that has been conceptualized as "moral injury" (Litz et al. 2009).

2.3 Test battery

The test battery is split in two parts. The first part will be applied before the simulation, i.e. intends to measure demographics, traits and state measure baselines as well as the heart rate baseline. The questionnaire includes basic demographic data (gender, age, ROTC involvement, year of study), video game experience, trait measures for altruism (Costa and McCrae, 1992), dutifulness (Conn and Rieke, 1994), war and peace attitude (adaptation of Van der Linden et al., 2011), perceived stress reactivity (Schlotz et al., 2011), and moral orientation (Narvaez et al., 2011). In addition, a state aggression measure (Spielberger, 1999) was used.

The second part intends to capture the influence of the simulation on moral reasoning. It first includes a "reason test" in which the participants will have to evaluate arguments that justify decisions in

Trolley like dilemmas (i.e. killing one for saving five persons) using four dimensions: Two dimensions capture a “cognitive component”, i.e. the participants evaluate on a 6-point Likert scale whether they consider a specific justification to be universalizable (applicable to everyone; end points: supporting this argument depends on your personal opinion vs. everybody should support this argument) and logical (end points: This argument doesn’t support the decision at all, vs. This argument clearly supports the decision.). Two dimensions capture an “emotional component”, i.e. the participants evaluate whether they find a specific justification appealing (end points: This argument does not at all feel right for me. vs. This argument feels completely right for me) and publicize (end points: I would not use this argument in public at all. vs. I certainly would use this argument in public.)

Second, it includes a value evaluation. Seven values out of an extended value study (Christen et al., in preparation) were chosen that are related to the problem under investigation: Aggressiveness, authority, cooperation, loyalty, obedience, responsibility, and security. The values are presented with representative descriptions and participants then evaluate the values along four dimensions: moral (the degree to which the value claims to be universally valid and corresponding actions are judged as right or wrong), official (the degree to which the value relates to compliance towards internal rules or laws), community oriented (the degree to which the value supports the cooperative being-together of people) and emotional appeal (the degree to which following this value feels right for the person).

In addition, the test battery included a memory and a perception test. In the memory test, the participants were asked to recall the sequence of the landmarks that characterize the five missions – these landmarks were both visible and mentioned in the prompts the participants receive. In the perception test, the participants were asked to guess the distance between the own position (“drone”) and the target (“persons”). Both measures are intended to capture the degree of “detactment” from the decisions, which may serve as an indirect measure of social stress. In the second part of the study the aggression state measure is taken again, whereas the difference between pre and post scenario serves as indirect measure for psychological stress.

In the semi-structured interview used for debriefing, the participant will report their experiences during the simulation and also indicate whether or not they consider each of the missions a success.

3 Results of Pre-Tests

We have conducted three pre-test studies for the test battery. The first two were used to determine suitable items for the “reason test”. The third pre-test served to check the sensitivity of the whole test battery and to obtain a “baseline” for the measures used such that we can check for systematic biases due to our sample in the main study. Participants were recruited using Amazon Mechanical Turk, which has been shown to be a superior subject pool compared to college undergraduates as the social diversity is higher (Buhrmester et al., 2011). Below, we briefly report on the main findings:

- In the first study (n=311) we identified 24 arguments out of 70 (justification for either choice in the Trolley dilemma) where the cognitive and emotional dimensions dissociate strong and where significant differences between people that have chosen to redirect versus people that have chosen not to redirect were found.
- In the second study (n=381) we repeated the first study only for the sub-set of 24 arguments identified above. Arguments were excluded where significant differences were found between study 1 and 2 (test-retest check) and where the dissociation between the cognitive and emotional dimensions was low. A total of 12 reasons (6 pairs) were then selected for the reason test.
- In the third study (n=204), the whole test battery as described in section 2.3 was pre-tested. We also included questions to identify persons with military experience (n=19) and persons that already experienced tragic dilemmas in their life (n=22). We then checked whether group

differences regarding our measures can be found. Significant differences were found with respect to aggression, altruism, some aspects of moral identity, and stress reactivity. Also the value test provided significant differences for several values, and the reason test showed significant differences in the emotion dimensions. Thus, the data indicates that our test battery is sensitive to the types of changes that result from experiencing different scenarios in our simulation.

Usability tests of the simulation with a total of ten participants demonstrated that they are able to operate the simulated RPAs and to cope with the orders given by the system. The whole experiment including instructions and debriefing takes about one hour. The main study (n=200) started in March and is ongoing.

4 Discussion and Conclusion

Human-computer interaction in the context of Remotely Piloted Aircrafts reveals effects that go beyond mere cognitive and executive functioning. In particular in combat, the moral dimension of what the RPA pilots are actually doing may be a significant factor for explaining the psychological distress those system operators experience. Therefore, the “detachment effect” that the “technological system RPA” is expected to generate due to a purely screen-based interaction with the mission and targets may actually not diminish the moral stress that RPA pilots experience. Although the remote-controlling of weaponized systems indeed protects the body of the pilot, the mind still may still be vulnerable. The lack of physiological involvement of the body in combat may undermine a mechanism that allows for justification of the actions, e.g. following the argument of “self-defense”. This then would require a more abstract justification that shows up in a higher agreement for – otherwise not supported – reasons that justify killings and a de-coupling of the emotional and cognitive evaluation dimensions of these reasons. If the main study indeed finds such changes in the evaluation of reasons and values associated with the dilemma and if these changes are bigger in the military scenario compared to the disaster management scenario, we will have found indications that the combination of remote-controlled killing and close but detached visual observation still have the potential to induce moral harm in the operator of these systems.

An interesting follow-up question is whether changes in the design of the human-RPA interaction may diminish the moral stress experienced by its operators. In the current study, the main focus is the interaction between the pilot and the simulated dilemma and not the interaction between pilot and the simulated drone. However, the set-up of our simulation will allow us to explore the actual man-machine interaction in follow-up studies, e.g. how changes in flight behavior relate to altered visual appearance of the scene. However, whether or not adapting the interface may be a way to mitigate the underlying moral problems that tragic dilemmas imply remains to be seen.

Acknowledgements: *This research has been supported by a grant of the Swiss Department of Defense, Armasuisse Research & Technology.*

References

Awad, A., Ghobashy, M., Ouda, W., Stout, R. Silverman, D. G., Shelley, K.H. (2001). Different responses of ear and finger pulse oximeter wave form to cold presser test. Anesthesia and Analgesia, 92, 1483-1486.

Bleske-Rechek, A., Nelson, L. A., Baker, J. P., Remiker, M. W., and Brandt, S. J. (2010). Moral decisions in the Trolley Problem: People save five over one unless the one is young, genetically related, or a romantic partner. Journal of Social, Evolutionary, and Cultural Psychology, 4(3), 115-127.

Buhrmester, M.D., Kwang, T., and Gosling, S.D. (2011). Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality, data? Perspectives on Psychological Science, 3 (6), 13-5.

Bumiller, E. (2011). Air Force Drone Operators Report High Levels of Stress. NY Times. http://www.nytimes.com/2011/12/19/world/asia/air-force-drone-operators-show-high-levels-of-stress.html (last access: December 8, 2013).

Christen, M., Narvaez, D., Tanner, C. (in preparation). A German-English comparison of synonymy-based value maps reveals underlying semantic structures of cultural moral differences.

Cohen, J. (2007). Drone spy plane helps fight California fires. Science 2, 318, 727.

Conn, S. R. and Rieke, M. L. (1994). The 16PF Fifth Edition technical manual. Champaign, IL: Institute for Personality and Ability Testing.

CORDIS (2004): Software and fire hose enhances fire-fighting resources. European Community Research and Development Information Service. Project reference: EVG1-CT-2000-30002. Available at: http://cordis.europa.eu/result/brief/rcn/1357_en.html (last access: December 8 2013)

Costa, P. T., Jr. and McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Cummings, M.L., and Guerlain, S. (2007). Developing operator capacity estimates for supervisory control of autonomous vehicles. Human Factors 49: 1–15.

Dillow, C. (2013). Now hiring drone jockeys. Fortune 167(1), 41 (January 14 2013). Available at: http://tech.fortune.cnn.com/2013/01/11/drone-pilots-jobs/ (last access: December 8 2013).

Dixon, S., Wickens, C.D., and Chang. D. (2005). Mission control of multiple unmanned aerial vehicles: A workload analysis. Human Factors 47, 479–487.

Donmez, B., Cummings, M. L., and Hudson G. D. (2009). Auditory decision aiding in supervisory control of multiple unmanned aerial vehicles. Human Factors, 51(5), pp. 718-729.

Enemark, C. (2013). Unmanned drones and the ethics of war. In Routledge Handbook of Ethics and War (Allhof, F., Evans, N. G., and Henschke, A., eds.). p. 327-337, Routledge, New York.

Foot, P. (1967). The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices. Oxford Review, 5, 5-15.

Gamer, M., Rill, H., Vossel, G., and Gödert, H. (2006). Psychophysiological and vocal measures in the detection of guilty knowledge. International Journal of Psychophysiology, 60 (1), 76-87.

Gamer, M., Verschuere, B., Crombez, G., and Vossel, G. (2008). Combining physiological measures in the detection of concealed information. Physiology & Behavior, 95, 333-340.

Gregory, D. (2011). From a view to a kill: drones and late modern war. Theory, Culture and Society, 28(7-8), 188-215.

Gunn, D. V., Warm, J. S., Nelson, W. T., Bolia, R. S., Schumsky, D. A., and Kevin J. Corcoran, K. J. (2005). Target acquisition with UAVs: Vigilance displays and advanced cuing interfaces. Human Factors: The Journal of the Human Factors and Ergonomics Society, 47, 488-497.

Kmietowicz, Z. (2012). Use of drones is a war crime and must be stopped, says charity. British Medical Journal, 345, e6820.

Lee, P. (2012). Remoteness, Risk and Air Crew Ethos. Air Power Review, 15(1), 1-20.

Lindlaw, S. (2008). UAV operators suffer war stress. Air ForceTimes, available at http://www.airforcetimes.com/article/20080807/NEWS/808070315/UAV-operators-suffer-war-stress (last access: December 8, 2013)

Litz, B.T., Stein, N., Delaney, E., Lebowitz, L., Nash, W.P., Silva, C., and Maguen, S. (2009). Moral injury and moral repair in war veterans: a preliminary model and intervention strategy. Clinical Psychology Review, 29(8): 695-706.

Lützén, K., Cronqvist, A., Magnusson, A., and Andersson, L. (2003). Moral Stress: synthesis of a concept. Nursing Ethics, 10, 312-322.

McCrisken, T. (2013). Obama's Drone War, Survival. Global Politics and Strategy, 55 (2), 97-122.

Medact.org (2013): Drones: the physical and psychological implications of a global theatre of war, update 2013. Available at: http://www.medact.org/wp-content/uploads/2013/11/report-dronesupdate-2013.pdf (last access: December 8, 2013)

Miller, G. (2012). Drone Wars. Science, 336, 842-843.

Mockenhaupt, B. (2009). We've seen the future, and it's unmanned. Esquire, November 2009. Available at: http://www.esquire.com/features/unmanned-aircraft-1109 (last accessed December 8, 2013).
Monat, A., and Lazarus, R.S. (1985). Stress and coping; an anthology. New York: Columbia University Press.
Narvaez, D., Brooks, J., and Mattan, B. (2011). Attachment-Related Variables Predict Moral Mindset and Moral Action. Society for Research in Child Development, Montreal.
Otto, J.L., Webber, B.J. (2013). Mental Health Diagnoses and Counseling Among Pilots of Remotely Piloted Aircraft in the United States Air Force. Medical Surveillance Monthly Report, 20(3), 3-8.
Ortega, H. (2013). Challenges in Monitoring and Maintaining the Health of Pilots Engaged in Telewarfare. Medical Surveillance Monthly Report, 20(3), 2.
Pyszczynski, T., Solomon, S., and Greenberg, J. (2003). In the wake of 9/11: The psychology of terror. American Psychological Association, Washington, DC.
Sauer, F., and Schörnig, N. (2012). Killer Drones—the silver bullet of democratic warfare? Security Dialogue, 43(4), 353–370.
Schlotz, W., Yim, I.S., Zoccola, P.M., Jansen. L., and Schulz, P. (2011). The Perceived Stress Reactivity Scale: Measurement invariance, stability, and validity in three countries. Psychological Assessment, 23(1), 80-94.
Schneiderman, R. (2012). Unmanned Drones Are Flying High in the Military/Aerospace Sector. IEEE Signal Processing Magazine, 29(1), 8-11.
Shaw, I. G. R., and Akhter, M. (2012). The Unbearable Humanness of Drone Warfare in FATA, Pakistan. Antipode, 44(4), 1490-1509.
Shelley, K. (2007). Photoplethysmography: Beyond the calculation of arterial oxygen saturation and heart rate. Anesthesia and Analgesia, 105, S31-S36.
Singer, P. (2009). Wired for War. Penguin Press, New York.
Spielberger, C. D. (1999). Professional Manual for the State-Trait Anger Expression Inventory (STAX-2). Odessa, FL: Psychological Assessment Resources.
Thomson, J. J. (1985). The Trolley Problem. The Yale Law Journal, 94(6), 1395-1415.
Trimble, S. (2008). Flying Predators Bad For Pilot's Health The DEW Line, 17 April 2008, Flightglobal. Available at: http://www.flightglobal.com/blogs/the-dewline/2008/04/study-flying-predators-bad-for/ (last access: December 8, 2013).
USAF (United States Air Force). (2009). Factsheet MQ-1 Predator Unmanned Aerial Vehicle. Available at: http://www.162fw.ang.af.mil/resources/factsheets/factsheet.asp?id=11932 (last access: December 8, 2013).
Valve (2013) Source Information sheet. Available at: http://www.valvesoftware.com/SOURCE_InfoSheet.pdf (last access: December 8, 2013).
Van der Linden, N., Bizumic, B., Stubager, R., and Mellon, S. (2011). Social representational correlates of attitudes toward peace and war: A cross-cultural analysis in the United States and Denmark. Peace and Conflict: Journal of Peace Psychology, 17 (3), 217-242.
Wilkinson, J.M. (1987/88). Moral distress in nursing practice: experience and effect. Nursing Forum, 23(1), 16-29.