Reviewer #1: The paper presents an algorithm for task allocation and path planning for first responders (FRs) in the aftermath of major disasters.  The authors use the AtomicOrchid platform to compare the behavior and performance of people playing a real-world mixed reality game with and without the assistance of the task allocation and path planning agent.  The analysis of the results indicate that the task allocation made by the agent improved people's performance.

While I do not see much contribution in the path planning algorithm per-se (the algorithm is naive and does not consider human response to the advised path), I do believe the task allocation algorithm is of much importance.  Specifically, the dynamic nature of the algorithm in reallocating tasks if the originals are rejected by the human first responders and the ability of the algorithm to give a real time allocation due to its approximate solution to MMDP.

The paper is very relevant to multiagent systems since it presents a recommendation agent that is able to cooperate with people and dynamically adjust to their actions.

The paper generally compares the current algorithm to previous algorithms but not empirically.  Still, the idea is mostly to provide a proof of concept for an algorithm that is able to cope with people's unpredicted actions and hence I don't think the empirical comparison to previous algorithms is a must.

The paper is well written.  The introduction and related work sections are thorough and supply extensive review of related work.  Still, there are several issues that should be seriously considered:

1. In the experimental design section the paper does not mention if there was an incentive for the participants to act the same way they act in real life.  If there was no incentive I do not understand the benefit of experimenting in a "real-world mixed-reality game".  In order to evaluate a system/algorithm in the real world people must have an incentive to act as they act in a real world scenario.

***RESPONSE****: The participants were told that their goal was to save as many targets as possible and the participants were seen to engage with the game as they often ran to save targets before they were covered by the radioactive cloud. Thus the participants clearly immersed themselves into the game and enjoyed the fun-factor associated. The benefit of the real-world mixed reality game is that it helps us uncover the challenges faced by participants in reacting to agent-based instructions in conditions of time pressure, stress, and physical fatigue. This is not possible in computer based simulations or lab studies****.*** *We clarify this point with sentences in Sections 6.1 and 6.4.*

2. In the results section the paper does not mention any statistical significance test that were performed.  The experiments were conducted on a relatively small group of participants (8 for each treatment), In order to convince the reader that indeed the treatment (agent guided) improved people's performance, statistical significance tests should be provided. Needless to say that a drastic change in the performance of only one participant, when relying on 8 participants, can substantially change the conclusions.

***RESPONSE****: we clearly stated* ***(section 6.1)*** *we are not focused on statistically significant results but use the numbers to guide the qualitative evaluation of the experiments. This approach is common in HCI studies and this is what we adopted here. We analysed more than 20 hours of videos (of participants/HQ), cross matched events in the system with events in the video recordings, and analysed all the messages involved in order to generate meaningful conclusions.*

3. The experiment was conducted on one specific scenario and configuration (the target locations, pattern of cloud movement, number of targets and starting points).  I understand the difficulties in recruiting people for real-world experiments, however it is impossible to generalize based on one specific configuration.

***RESPONSE****: we agree it is impossible to generalize based on one specific configuration. This is not the point of the evaluation. The approach stems from contextual analyses of interactions (see Suchman’s work) that aim to uncover interactional issues that arise in a specific setting and for a specific set of users. Such experiments will generate different results in different contexts but the learning derived from such runs does let us develop some design guidelines do address those issues. This is reinforced in 6.1.*

4. The difference in performance between the groups that played the game with the agent and the group that played the game without the agent can be explained by the quality of the HQ (the person that gave the advice).  In order to determine that the difference is indeed due to the agent further experiments are required.

***RESPONSE****: We disagree that further experiments are required. In all experiments, participants knowledgeable of the task allocation problem – recruited from a research team - manned the HQ. We observed the HQ performance using video analysis and also interviewed them afterwards. It became clear that the HQ operator was overwhelmed with tasks in the non-agent case and underperformed as a result. In the agent-based case, the HQ specifically relied on the agent to do the allocation. This we see as a strength of the approach. We have made this clearer (in Section 6.4) and included statements from the participants at the end of Section 6.4 to reflect these behaviours.*

Suggested improvements:

1.  As discussed above, the paper can be more convincing if more people would have been recruited for the experiments.  I would like to see more field trials and if possible to add another world configuration for the experiments.

**RESPONSE**: *as per normal HCI experiements, the number of experiments and people involved is perfectly valid (see 6.1).*

2. In the algorithm for allocating tasks there is no reference to the probability that this task will be accepted by real people and there is no reference to the expected health level after completing the task.  I believe these two points should be included in the algorithm to reduce the rejection level.

**RESPONSE***: In our algorithm, the expected team value is a reference to the expected health levels after completing the task. Note that the expected team value is estimated by several simulations. During each simulation, we simulate the changing of the FRs’ health level when they are in the radioactive cloud. If a FR’s health level is below 0, a penalty is returned for estimating the team values. To simplify our planning process, we did not consider the probability that a task will be rejected by FRs. If a plan is rejected, we simply rerun the planning process and generate a new plan. We added more explanations of this procedure in Section 4.*

3. In figure 4, it is not clear what is the meaning of the "overridden" box.

**RESPONSE** *We mentioned that H did override the instructions by PA in some cases – this is what the box refers to (1st para page 21). We further added a note in the caption to Figure 5 to explain this.*

4. The agent's algorithm is divided into two; the first part is to dynamically allocate tasks to people and the second part to compute the best path.  In the results section there is no reference to people's deviation from the path offered to them.  Did the FRs understand the path planner? Did they choose a different path than the one advised to them?

**RESPONSE**:

*The path planner computes the shortest path (a set of contiguous cells in the map, considering potential obstacles). Participants were sometimes disorientated and took other paths than those suggested by the agent. However, in some cases, even if they took longer paths, they walked faster (or ran) to complete tasks in shorter times than expected. We have added examples at the end of 6.4 to clarify this.*

5. Once a task was accepted by a FR, can the FR regret and abandon the task? If so, how would the agent be informed?

**RESPONSE:** *If a FR rejects and abandons the previously accepted task, a new plan is generated for him/her by the agent based on the current state and the preference of the FR until they accept it. Once they accept, the agent assumes they will complete it. This is clarified in Section 6.*

6. In page 8 line 43 the notation is a bit confusing, I believe the authors meant: O\_t\_j subset of {O\_i | P\_i in C\_j}.

RESPONSE:*Thank you. This has been fixed.*

7. Page 8 line 47: "for the commander and to find".

**RESPONSE: Fixed**

8.  Page 11 line 40: "Equation 2"

**RESPONSE:** *This is correct*

9. Please check carefully the references, for example, reference #55 does not have the full name of the paper.

**RESPONSE:** *Thank you. Fixed and checked others.*

Reviewer #3: JOURNAL OF AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS REVIEW FORM:

SUMMARY. PROVIDE A BRIEF SUMMARY OF THE REASONS FOR YOUR RECOMMENDATION. WHAT ARE THE MAJOR STRENGTHS AND WEAKNESSES OF THE MANUSCRIPT?

This paper presents an approach to using agents and agent technology assist humans in a case of disaster response.  Algorithms for doing a task allocation and path planning under uncertainty are presented, but the focus and emphasis is on running the system with real humans in a "realistic" scenario.

RELEVANCE. HOW RELEVANT IS THE PAPER TO THE DISCIPLINES OF AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS?

The work falls clearly into the field of multiagent systems and aims to push the technology into thinking about humans and agents cooperating together to achieve something.

SIGNIFICANCE [NOT APPLICABLE TO SURVEY PAPERS]. HOW SIGNIFICANTLY DOES THE PAPER ADVANCE THE CURRENT STATE-OF-THE-ART? HOW BIG AN IMPACT WILL THE PAPER HAVE, ON WHAT AREAS, AND WHY?  WILL THE PAPER STIMULATE FURTHER RESEARCH?

I think the paper over-states its general significance and it is more a significant step towards the author's own efforts than something general, i.e., it represents a significant step towards building a nice platform but the results are not necessarily of general significance.

**RESPONSE**: *We disagree that the efforts are mainly towards building a platform –This paper evaluates an interactional arrangement where the agent bypasses the human operator at HQ. The results are the first of their kind in terms of having a planner agent directing humans in completing real-world tasks.*

I think the discussion of previous work is overly narrow and is not supported by the particular work.  First, there are examples of other work using real humans in disaster response experiments with agents, e.g., <http://www.computer.org/csdl/proceedings/saint/2007/2756/00/27560002.pdf> and Wirz, Martin, Daniel Roggen, and G. Troster. "User acceptance study of a mobile system for assistance during emergency situations at large-scale events." Human-Centric Computing (HumanCom), 2010 3rd International Conference on. IEEE, 2010.  These may not be precisely what was done here, but seem to be relevant.

**RESPONSE**:

*We have included the suggested references. We make it clear that our work is focused on ‘Coordination’ that involves humans pairing up to complete tasks. The works mentioned point to ‘evacuation’ which does not involve teaming up humans to perform specific tasks. We have clarified this point in the related work (Section 2.2) and add some pointers to the non-coordination work.*

Moreover, more generally there has been a lot of work with humans and agents, with humans directing agents.  It might be argued that no one has done the type of analysis here or been in the domain of disaster response, but I don't think that that makes the work not relevant.   For example, Tambe's security work has game theoretic algorithms telling security services which routes to patrol.  On a larger, and less academic, scale organizations such as FedEx have been using agent-style algorithms to tell drivers where and how to go for many years.  Newer systems, like Uber, even do this in an almost distributed way for taxi-drivers.  The argument that this work is not relevant because of the disaster response domain doesn't really resonant because the subjects in this work are actually just students with 30 minutes training, so no more first responders than FedEx drivers are.  I think if the authors want to claim there is a hole in the literature, they need to do it much more

carefully.

**RESPONSE**: *the reviewer is not completely correct about the nature of our work and its relationship to others and we agree this makes our work sound less innovative. Specifically, we have clarified (in Section 2) that:*

1. *Tambe’s work is not about real-time coordination. Their work pre-computes plans for security guards.*
2. *Our work aims to study the challenges in human-agent collaboration – not simply agent-based guidance.*

*Hence, we now delineate our work from these real-world examples of agent-guided human systems more carefully in our related work.*

The contribution of the paper is made clear but is not a crisp contribution that is easy to evaluate.  There is a "novel" algorithm that is only briefly compared to some simple approaches in an Appendix.  The field trials are promising but so small and noisy that it is hard to draw conclusions.  The contribution is aimed at being the combination of these things and it is not clear what the generality of that is.

RESPONSE: *We have expanded some of the discussions on the algorithm and further explained how we delineate from previous work on the algorithmic side. We also expanded on the discussion of the results (as in response to Reviewer 1).*

There are three major conclusions drawn in the main paper, under "In summary, these results suggest three key observations with regard to human-agent coordination in the trial:"  I think the first parts of the first two can be justified from the data, but then the authors add discussion that is not supported by the data.  "may reflect the better forward-planning capability of the planning agent compared to human FRs" - must at least come with caveats about the specific problem, since it is likely that in a real, unfolding disaster humans do much better at keeping resources in reserve, devising high-level strategies in the face of restricted resources and so on, things that the MMDP does not even model.

**RESPONSE**: We have added this caveat to the bullet point (page 23) in response to the reviewer’s comment.

  "Without this facility we believe the FRs would have chosen to ignore the plan." - maybe, but its also possible they only reject because they know they can reject, but if there was no option to reject they wouldn't just do whatever they liked, they might not be happy, but

they'd treat the agent like their boss - who knows, nothing in the data tells us one way or the other.

**RESPONSE**: We now explain how this conclusion was drawn from the results of the field trials and prior pilot studies.

 The third conclusion, ". When task allocation was handled by P A, H could focus on providing vital situational awareness to safely route first FRs around danger zones: thus demonstrating effective division of labour and complementary collaboration between humans and agents." may be correct, may have been what was (objectively) observed by the authors but there is no data presented to support it.  The authors are careful and introduce the conclusions with the words "suggest" and "observations" rather than anything stronger - but if these are not the conclusions then there are none.

There are additional conclusions in the "Conclusion" section, outside of the evaluation section.  These are broad observations and may be correct but are not strongly supported by data.  Nor are they particularly non-intuitive.

**RESPONSE**: *We have carefully analysed a lot of video data and chat logs (and added explanations of these in Section 6.1) –these are standard approaches in such field trials. We have now also added statements from the participants to support the conclusions we draw.*

The MMDP solution is clearly not presented in enough detail to allow replication, but for the purposes of this paper is probably OK.

**RESPONSE**: *We built a MMDP simulator of our scenario for planning. In more detail, the factors related to the plan including the radioactive level of each location, the location of each FR, their health level, and the status of each task are represented as state variables. The actions of each FR are: staying in the current location, moving to a neighbour location, or completing a task at the current location. The transition function defines the probabilities of transiting from the current state to a next state given a joint action of all FRs. The FRs are rewarded if a task is completed before its deadline and get a penalty if they are killed by the radioactive cloud. The policy is a mapping from a state to a joint action for all FRs so each FR can select an action for the current state. This is clarified in Section 4.*

FIGURES AND TABLES.  ARE THE FIGURES AND TABLES NECESSARY AND SUFFICIENT?

For this sort of work, where things are actually moving around an environment, I think it is always useful to show some sort of trace of movement (it should exist since the agent knows where the humans are?) to give a feel for what happened.  Even just an analysis of movement would be useful. \*Why\* were the unsupported humans worse, did they keep criss-crossing the environment or were they myopic and missed opportunities or did they just interfere/conflict with each other? - any of these cases suggests something very different for why the agent support was useful.

**RESPONSE**: *We have added a figure (Figure 4) to show a trace of how users moved in the environment. We added a discussion of patterns observed.*

This paper is interesting and reflects quality work.  However, it is just not clear what the value of it is to the broader community.  It goes shallowly into multiple areas, making it hard to draw general conclusions on any of the areas and not clear what to do with the high-level conclusion.

**RESPONSE:** *we provide a deeper analysis and positioning of the work in the new version in our related work (Section 2 – especially in response to Reviewer2) and provided more details on the experiments (Section 6).*

- Better situation within existing work: It is entirely unclear why extensive previous work on humans interacting with intelligent systems in a range of domains does not apply here (especially when students are standing in for first responders)

- Clearer conclusions that are better supported by data.

**RESPONSE**: *SEE responses above as to changes made*.

- Collect more data.

- Try different "agent" algorithms, it is completely unclear whether some trivial agent algorithm that just made sure humans were going to a nearby task or not have multiple humans go to the same task wouldn't achieve mostly the same results.

**RESPONSE***: We do not believe running more experiments will generate more intuitions (we have done so but not found new insights that differ from what we report here). Our simulations confirm that a simple algorithm would underperform. Hence, we do not believe it is worth spending a lot of effort using an obviously worse benchmark.*

Reviewer #4: The objective of this article is to study human-agent collaboration for disaster response when taking into consideration uncertainties in environments and FRs' responses to plans generated by agents. Based on my reading and understanding, I find the materials presented in this manuscript interesting and the approach is technically sound. Here are a few suggestions for the possible improvement.

(a) It appears that the abstract is rather long. The authors are suggested to make it more concise.

**RESPONSE:** *We shortened the abstract as far as possible.*

(b) The results presented in Section 6 are very interesting. For better understanding the possible response of FRs. It might be a good idea to test both light task load and heavy task load. In other words, FRs should respond differently given light/heavy task load. Then different strategies are expected to deal with these two cases. Essentially, PA can predict the possible response from FRs and take the prediction into consideration when assigning tasks.

**RESPONSE**: *the workload was already perceived as being quite heavy in the experiments. We have added a description of discussions in the debriefing session (Section 6.4) to exemplify what workload the H participants faced in the agent/non-agent scenarios.*

(c) How to pick the value v(C\_jk) in (2)? Since it has a direct impact on the optimal solution, better to give more clues on how to pick this value.

**RESPONSE:**: *The estimate of v(C\_jk) is detailed in Section 4.1.1. Specifically, the team value v(C\_jk) is estimated by several simulation as shown in Algorithm 2. We now link Equation (2) to Section 4.1.1 in the text.*