Adaptive Task Allocation Preferences in Different Workload Scenarios in Driving Automation Systems

Skye Taylor[1], Katherine Garcia[2], Jing Chen[2*], Bin Hu[2]

[1]Virginia Tech; [2]Old Dominion University

Adaptive task allocation is used in many human-machine systems and has been proven to improve operators' performance with automated systems. However, there has been limited knowledge surrounding the benefits of adaptive task allocation in automated vehicles. In this study, participants were presented with photos and videos depicting driving scenarios of low or high workloads at two levels of automation (SAE Levels 2 and 3). The participants reported which tasks they felt comfortable allocating to themselves or to the driving automation system (DAS) in each driving scenario, as well as whether they would conduct the task allocation manually or have the DAS automatically allocate the tasks. Our results showed that participants preferred conducting manual task allocation and preferred the system to complete more tasks when the perceived workload was high. There was no significant difference between the high and low workload scenarios in terms of whether participants chose to allocate tasks.

While driving automation system (DAS) are receiving more attention, developers are far from providing the market with fully "self-driving" cars, such as SAE Level 5 vehicles where driver input or monitoring is not required by the system (SAE, 2021). SAE Levels 2 and 3 require users to monitor the system to complete driving tasks and are currently the highest-level DAS on the market (SAE, 2021). The human driver is still a necessary component in these DASs for successful operation of the vehicle. However, the question remains as what tasks the human should step in and take control over the DAS for and when.

Adaptive task allocation (ATA) is a contending solution to the distribution of driving responsibilities between the system and the human operator. ATA is defined as, "in ergonomics, a system design that supports allocation of tasks to the human operator or the machine according to the state of the system, the state of the operator, or other operational rules" (American Psychological Association, n.d., Adaptive Task Allocation section). ATA has been shown to improve the monitoring of many automated systems such as aerial vehicles (Parasuraman et al., 1996), human-robot teams (Dubois & Ny, 2020), and automated supervisory control systems (Gutzwiller et al., 2015). Yet, ATA has not been widely researched in the specific application domain of DASs.

With automated systems, task allocation can be completed manually by the human operator (Klien et al., 2004), or automatically by the system (Parasuraman et al., 2000). In the DAS domain, manual task allocation is where the human driver delegates a task to be completed by themselves or the vehicle (e.g., assigning a lane-changing task to themselves), whereas automatic task allocation is conducted by the automation to delegate a task to the driver or the vehicle without direct intervention from the driver (e.g., assigning the steering task to the vehicle; Klien et al., 2004; Parasuraman et al., 2000).

When completing dual tasks like driving, higher workloads result in poorer performance (Moacdieh et al., 2020). While many researchers commend adaptive techniques for purposes of trust in automation (Inagaki, 2003; Parasuraman et al., 1992), these techniques require the user to work with the system and monitor its behaviors. It is evident that monitoring automated systems impose significant workload (Kaber & Endsley, 2004).

This study examined the task-allocation preferences in different driving scenarios across SAE Levels 2 and 3 DASs (SAE, 2021). Level 2 DASs requires the drivers' constant supervision of the vehicle, whereas Level 3 only requires the drivers' attention when prompted to takeover (SAE, 2021). This distinction is expected to create a difference in the preference of task assignment (Marinik et al., 2014). Participants were provided with a description of the capabilities of their "vehicle", a series of driving scenarios of different workload levels, and were asked about their preference regarding their preferred method of task allocation and the tasks (primary driving tasks and secondary tasks) they would assign to themselves or the system under the conditions presented.

It was hypothesized that participants would prefer manual allocation over automated task allocation at higher levels of automation (LOAs) since this behavior has been shown in other automated systems (Klien et al., 2004). Additionally, it was expected that participants would be more likely to conduct manual task allocation at the higher LOA than the lower level since the system was more capable of conducting driving tasks while the manual allocation task was performed. Finally, due to higher capabilities at higher levels of automation (Hopkins & Schwanen, 2021), it was expected that primary driving tasks would be assigned to the higher-level vehicle (Level 3) more often than to the lower-level vehicle (Level 2).

Regarding the workload of the driving scenarios, it was hypothesized that participants would allocate a greater number of tasks to the vehicle when they encountered driving environments with a higher workload (Lyu et al., 2017). It was also expected for participants to prefer conducting manual task allocation in scenarios where the workload is low compared to when it is high since drivers are more likely to have a greater mental capacity at these times and that a high mental workload can lead to a decrease in driving speed to accommodate for a decrease in mental capacity (Landsdown et al., 2004).

METHOD

Participants

Participants were recruited via Amazon Mechanical Turk (MTurk), an online participant pool provided by Amazon. Eligible participants were required to be within the United States, have a valid driver's license, and be between 18 and 89 years of age. A total of 156 participants took part in the survey (102 males and 54 females). The ages of the participants ranged from 21 to 67 years, with a mean age of 36.29 years (*SD* = 8.34). Each participant was compensated $3.00 for completing the study which took an average of approximately 40 minutes.

Experimental Design

A 2×2 mixed factorial design was used, with LOA condition (low, high) manipulated between-subjects, and scenario workload (low, high) manipulate within-subjects. The SAE LOAs were used to distinguish between "High" automation (Level 3), and "Low" automation (Level 2; SAE, 2021). Photos and videos of different driving scenarios with varying locations and driving conditions were used to manipulate the workload level within-subjects. For example, a low workload driving scenario depicted an open road with minimal visual stimuli, such as very few pedestrians, while a high workload scenario depicted a busy road with multiple vehicles, signs, and poor weather conditions. Each participant experienced an equal number of "high" and "low" workload driving scenarios presented randomly throughout the study.

The dependent variables included the percentage of times participants chose to conduct the task allocation, and the percentage of times participants preferred to use manual task allocation rather than automatic task allocation. The perceived workload of the depicted driving scenarios measured by the NASA TLX (Hart & Staveland, 1988) were also collected, using a 20-point scale in the categories of performance, effort, frustration, mental, physical, and temporal demand. These six categories were averaged for an overall workload score.

Materials

This experiment was conducted online, hosted by Qualtrics XM. The study was comprised of 10 photos and 10 videos (each 30-seconds long) depicting different driving scenarios taken from the perspective of the driver. In each section, half of the photos/videos were intended to depict a "high workload" driving scenario (see Figure 1) with the remaining intended to depict a "low workload" driving scenario (see Figure 2). The materials were found on a website called Creative Commons, which provides materials for public use. In both the photo and video sections, participants were asked to conduct a task-allocation assignment for each driving scenario for five primary driving tasks and seven secondary driving tasks (see Table 1). They had to select whether they were more likely to complete each task themselves (*Assign to Self)*, have the vehicle complete the task automatically (*Assign to System)*, not complete the task (*Neither)*, or having no preference as to how the task was completed (*Neutral*). Additionally, the NASA TLX was used to measure the perceived workload for each depicted scenario on a 20-point scale (Hart & Staveland, 1988). Additionally, in the video section, participants were asked about their preference for when and how to conduct the task allocation. Participants answered either "Yes" or "No" for "Would you conduct a task allocation during the time period shown in the figure/video?", and "Manually" or "Automatically" for "Which of the following would you prefer during the time period depicted shown in the figure/video?".

Figure 1
Example of High Workload Depiction

Note. "Skier walking across the slushy street carrying skis, snowy overcast day, intersection, car, windshield wipers, red light, Wallingford, Seattle, Washington, USA" by Wonderlane is licensed under CC BY 2.0

Figure 2
Example of Low Workload Depiction

Note. "Open Road" by FotoSleuth is licensed under CC BY 2.0

Table 1
Summary of Tasks to be Allocated

Primary Driving Tasks	Secondary Driving Tasks
Steering	Operating windshield wipers
Braking	Operating turn signals
Accelerating	Operating infotainment system
Lane changing	Operating AC/heat
Monitoring surroundings	Activating headlights/high beams
	Activating/adjusting cruise control
	Sending text messages/making phone calls

Procedure

Upon beginning the study, participants were asked to review the project description and consent form. If the participants consented to participate in the study, they were prompted to answer some demographic questions. Participants were disqualified and unable to complete the survey if they did not meet the criteria for eligibility.

Next, the participants were provided with directions for completing the task allocation assignment and a description of the vehicle they would be "driving". If they were assigned to the Level 2 DAS, they were told that they would have to constantly monitor the roadway and take over at any time the vehicle was not functioning at full capacity. Participants assigned to drive a Level 3 DAS were told that the vehicle provides conditional automation and take over was only required when requested by the vehicle. Then, they were asked clarifying questions about the instructions and capabilities of their vehicle to confirm that they understood the capabilities of their vehicle and directions for the experiment. Participants were unable to proceed to the experiment until they answered all the clarifying questions correctly.

The online study included two main sections. The first section contained 10 photographs, and the second section contained 10 videos. Within each section, the five low workload scenarios and the five high workload ones were presented in a random order. For each photo/video the participants were asked to complete the task-allocation assignment for the driving-associated tasks, and the NASA TLX. The presentation order of the tasks in the task-allocation assignment was randomized. Additionally, for the video section, participants were asked about their preference on manual versus automatic task allocation, as well as whether they would conduct the task allocation in the scenario depicted.

After the last video clip and set of questions, participants were debriefed, thanked for their time, and given a unique code for MTurk to receive compensation for their participation.

RESULTS

Workload

A 2×2 mixed factorial analysis of variance (ANOVA) was conducted on the NASA TLX results for the 10 photos yielded a significant main effect of scenario workload $F(1, 154) = 7.29$, $p = .008$, $\eta_p^2 = .05$. The overall ratings were significantly higher for the photos corresponding with the "high" workload condition ($M = 11.53$, $SD = 4.16$) than the photos corresponding with the "low" workload condition ($M = 11.15$, $SD = 4.42$). The effect of LOA and interaction were not significant, Fs < 1.

A 2×2 mixed factorial ANOVA was also conducted on the NASA TLX results for the 10 videos. Frustration was the only component of the NASA TLX results for the videos section with statistical significance main effect of workload, $F(1, 154) = 9.58$, $p = .002$, $\eta_p^2 = .06$. The ratings for frustration were significantly higher when the workload condition was "low" ($M = 9.72$, $SD = 5.70$) than when the workload condition was "high" ($M = 9.12$, $SD = 5.41$). The main effect of LOA was not significant, $F < 1$, so was the interaction between workload and LOA, $p > .05$.

Task Allocation Method

A 2x2 mixed factorial ANOVA was conducted on the percentage of the time participants answered "Yes" to conduct a task allocation during the time period shown for the 10 videos. Neither main effects of LOA nor scenario workload, nor their interaction effect were significant, Fs < 1.

A 2×2 mixed factorial ANOVA was also conducted on the percentage of time participants answered "Manually" to their preferred type of task allocation (manual vs. automated). A significant main effect of scenario workload was found, $F(1, 154) = 11.00$, $p = .001$, $\eta_p^2 = .07$. The average percentage of the time participants chose to conduct the task allocation manually was significantly higher when the workload condition was "high" ($M = 74\%$ $SD = .28$) than when the workload condition was "low" ($M = 67\%$, $SD = .30$). The main effect of LOA and interaction were not significant, Fs < 1.

Task Allocation Preference

For the photos, a 2×2 mixed factorial ANOVA on the percentage of time the option "Assign to System" was selected for each task, yielded a significant main effect of scenario workload for tasks overall, $F(1, 154) = 14.04$, $p < .001$, $\eta_p^2 = .08$. The average percentage of the time participants chose to assign the task to the system was greater when the workload condition was "low" ($M = 35\%$, $SD = .20$), than when the workload condition was "high" ($M = 30\%$, $SD = .17$). Specifically, this difference was significant for the tasks of "lane changing", "operating windshield wipers", "operating turn signals", "activating headlights/high beams", and "activating/adjusting cruise control" (see Table 2 and Figure 3). The main effect of LOA and interaction were not significant, $p > .05$.

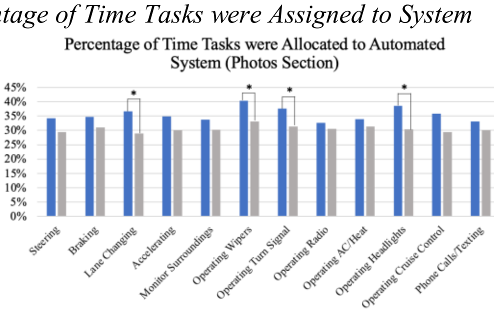
Table 2

Summary of Task Preference Results for Photos Section

Task	$F(1,154)$	p	η_p^2	Workload	M	SD
Lane changing	7.92	0.006	0.05	Low	36%	0.28
				High	29%	0.27
Operating windshield wipers	7.60	0.007	0.05	Low	40%	0.30
				High	33%	0.27
Operating turn signals	5.92	0.016	0.04	Low	38%	0.30
				High	31%	0.27
Activating headlights/ high beams	10.82	0.001	0.07	Low	38%	0.30
				High	30%	0.26
Activating/ adjusting cruise control	5.60	0.019	0.04	Low	36%	0.30
				High	30%	0.25

Figure 3

Percentage of Time Tasks were Assigned to System



Note. * Denotes a significant difference (α = 0.05)

For the videos, a 2×2 mixed factorial ANOVA was also conducted on the percentage of time the option “Assign to System” was selected for each task. The main effect of scenario workload was significant, $F(1, 154) = 9.80$, $p = .002$, $\eta_p^2 = .06$. The average percentage of the time participants chose to assign the task to the system was greater when the workload condition was “high” (M = 34%, SD = .20), than when the workload condition was “low” (M = 30%, SD = .18). Specifically, this difference was significant for the tasks of “lane changing”, “operating turn signals”, and “activating headlights/high beams”. The average percentage of the time participants chose to assign the task to the system was greater when the workload condition was “high” than when the workload condition was “low” for the operating turn signals and lane changing tasks. However, the opposite was true for the headlights task (See Table 3 and Figure 4).

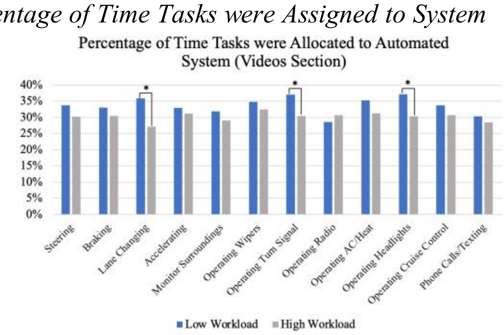
Table 3

Summary of Task Preference Results for Videos Section

Task	*F(1,154)*	*p*	η_p^2	Workload	*M*	*SD*
Lane changing	12.28	0.001	0.07	Low	36%	0.29
				High	27%	0.26
Operating turn signals	7.16	0.008	0.04	Low	37%	0.29
				High	30%	0.27
Activating headlights/ high beams	8.12	0.005	0.05	Low	37%	0.30
				High	30%	0.28

Figure 4

Percentage of Time Tasks were Assigned to System



Note. * Denotes a significant difference (α = 0.05)

DISCUSSION

The purpose of this research was to understand human drivers’ task-allocation preferences in DASs. An online study was conducted by presenting various driving scenarios to participants and inquiring about their preferences of task allocation (to themselves vs. the DAS) and task-allocation methods (manually vs. automatically), given different workloads and LOAs.

Workload

The results of the NASA TLX showed that our intended workload levels matched participants’ perception in the photos section; the perceived workload was significantly lower for the photos that were intended to depict a low workload scenario and significantly higher for those that were intended to depict a high workload scenario. This means that our manipulation of workload was valid for the photos. However, participants reported workload did not differ between our intended videos with low and high workloads, except that the frustration measure was greater when the workload condition was “low”. The phenomenon of “dissociation of workload and task demand” has been observed in several previous studies and could explain why the intended workload was not perceived by the participants (Lei et al., 2017, p. 95). Dissociation of workload and task demand means that the perceived workload was not linearly correlated with the actual demand of the task. This dissociation is often caused by “automaticity,” the result of completing a task repeatedly (Vidulich & Wickens, 1986). Frustration was likely higher for low workload scenarios because the participants were simply frustrated with having to watch a less stimulating video or because the intended workload of the environments presented were not well-perceived.

Task Allocation Method

It was expected that participants would prefer to conduct manual task allocation when the scenario depicted “low” workload (Landsdown et al., 2004). However, our results showed the opposite, that high workload driving environments elicit a greater preference for performing the task allocation manually. This finding could be because when drivers are in higher workload driving scenarios, they desire more control over how tasks are completed rather than trusting the vehicle to assign tasks (Dubois & Ny, 2020). Because DASs are a new and developing technology, drivers may not trust them to complete the challenging duty of allocating specific tasks (Parasuraman & Riley, 1997). Drivers are likely to trust their own abilities over the system’s, despite the description provided of the vehicle’s capabilities, and thus prefer to allocate tasks on their own.

Task Allocation

It was hypothesized that when driving in high workload environments, drivers would allocate more tasks to the vehicle. This was supported by results from the photo block, as tasks were more frequently assigned to the vehicle under high workload conditions. While the video stimuli provided less significant results, the trend showed that under higher scenario workload conditions, tasks were more likely to be assigned to the vehicle. This was true for all tasks except “operating headlights”. This may be because it is a very simple task that the operator feels they can manage despite the high workload. Another possible reason is that “operating headlights” may be considered a function that is based in communication with other humans, which explains why the driver prefers to complete this task rather than assigning it to the vehicle system.

Limitations and Future Studies

There are some limitations of this study due to the medium used. No significant findings were shown with regards to the LOA of the vehicle. This is likely because the manipulation of these conditions was simply a verbal description of the vehicle's capabilities. Additionally, since the interface allowed participants to manually allocate tasks to either themselves or the system, the LOA was not properly represented, because the more tasks allocated to the vehicle increase the level of the vehicles' automation (Hoffman and Johnson, 2019). The differences between Level 2 and Level 3 DASs would likely be more evident if the study was conducted in a driving simulator, allowing participants to interact with the DASs in distinct manners. In addition, the current study inquired about participants' preferences of task allocation and methods but did not include questions about the reasons behind their choices.

Future research can investigate why drivers prefer to allocate certain tasks to themselves and others to the automation, which can be utilized to further inform the design of task allocation DASs. To enhance the reliability of these results in a real-world setting, the study could utilize the same manipulation of high and low workload driving scenarios developed in a simulator with the technological capabilities of Level 2 and Level 3 automation. To further increase the effects of the results, a Likert scale, rather than a binary response question, could be utilized to determine how likely participants are to conduct manual task allocations during specific scenarios.

CONCLUSIONS

This study showed that there is a greater preference for manual task allocation in DASs when drivers are in high-workload scenarios than in low-workload scenarios. However, when drivers are in high workload environments, they are more likely to assign more tasks to the system. The results from this study and related research suggest that users may not prefer automated task allocation, but trust DASs to complete and manage complex tasks. When implementing ATA in DASs users prefer to be in the loop for task delegation. Perhaps using this method for task allocation, over time, will create a more dynamic interaction with human users and DASs. With the increasing popularity of personal DAS's, these findings should be considered when developing the models and algorithms used in DASs. Additionally, the information found with regards to user's preference of what tasks they prefer to complete on their own provide valuable insight on how road infrastructure for automated vehicles should be adapted.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1949760 and 2007386.

* Corresponding author: jc207@rice.edu

References

American Psychological Association. (n.d.). Adaptive task allocation. In *APA dictionary of psychology*. Retrieved June 18, 2021, from https://dictionary.apa.org/adaptive-task-allocation

Dubois, C. & Le Ny, J. (2020). Adaptive task allocation in human-machine teams with trust and workload cognitive models. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, (pp. 3241-3246). IEEE. doi: 10.1109/SMC42975.2020.9283461.

Gutzwiller R.S., Lange D.S., Reeder J., Morris R.L., Rodas O. (2015) Human-computer collaboration in adaptive supervisory control and function allocation of autonomous system teams. In *International Conference on Virtual, Augmented and Mixed Reality* (pp. 447-456). Springer, Cham. doi:10.1007/978-3-319-21067-4_46

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology* (Vol. 52, pp. 139-183). North-Holland. doi:10.1016/S0166-4115(08)62386-9

Hoffman, R. R., & Johnson, M. (2019). The quest for alternatives to "levels of automation" and "task allocation". In M. Mouloua & P. A. Hancock (Eds.), *Human Performance in Automated and Autonomous Systems* (pp. 43-68). CRC Press.

Hopkins, D., & Schwanen, T. (2021). Talking about automated vehicles: What do levels of automation do?. *Technology in Society*, *64*, 101488. doi:10.1016/j.techsoc.2020.101488

Inagaki, T. (2003). Adaptive automation: Sharing and trading of control. In E. Hollnagel (Eds.), *Handbook of Cognitive Task Design* (pp. 147-169). CRC Press.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness, and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, *5*(2), 113–153. doi:10.1080/1463922021000054335

Klien, G., Woods, D., Bradshaw, J., Hoffman, R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems, 19*(6), 91–95. doi:10.1109/MIS.2004.74

Landsdown, T. C., Brook-Carter, N., & Kersloot, T. (2004). Distraction from multiple in-vehicle secondary tasks: Vehicle performance and mental workload implications. *Ergonomics, 47*(1), 91-104. doi:10.1080/00140130310001629775

Lei, S., Toriizuka, T., & Roetting, M. (2017). Driver adaptive task allocation: A field driving study. *Le Travail Humain,* 80(1), 93-112. doi:10.3917/th.801.0093

Lyu, N., Xie, L., Wu, C., Fu, Q., & Deng, C. (2017). Driver's cognitive workload and driving performance under traffic sign information exposure in complex environments: A case study of the highways in China. *International Journal of Environmental Research and Public Health*, *14*(2), 203. doi:10.3390/ijerph14020203

Marinik, A., Bishop, R., Fitchett, V., Morgan, J. F., Trimble, T. E., & Blanco, M. (2014). *Human factors evaluation of level 2 and level 3 automated driving concepts: Concepts of operation* (No. DOT HS 812 044). United States. National Highway Traffic Safety Administration.

Moacdieh, N. M., Devlin, S. P., Jundi, H., & Riggs, S. L. (2020). Effects of workload and workload transitions on attention allocation in a dual-task environment: Evidence from eye tracking metrics. *Journal of Cognitive Engineering and Decision Making*, *14*(2), 132-151. doi:10.1177/1555343419892184

Parasuraman, R., Mouloua, M., Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors, 38*(4), 665-679. doi:10.1518/001872096778827279

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*, 230-253. doi:10.1518/001872097778543886

Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems*. Catholic Univ of America Washington DC Cognitive Science Lab.

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Sysems, Man, and Cybernetics-Part A: Systems and Humans, 30*(3), 286-297. doi:10.1109/3468.844354

SAE International. (2021). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles* (Publication No. J3016_202104).

Vidulich, M. A., & Wickens, C. D. (1986). Causes of dissociation between subjective workload measures and performance: Caveats for the use of subjective assessments. *Applied Ergonomics*, *17*(4), 291–296. doi:10.1016/0003-6870(86)90132-8