Solving Complex Dynamic Tasks Through Feature Focus

Jana B. Jarecki

Pinja Haikka

Jacon Sherson

Solving Complex Dynamic Tasks Through Feature Focus

Humans can solve tasks that are computationally highly complex, such as image content classification, face recognition, or smooth visual motion tracking (REF, REF, REF). Recent work by Sørensen et al. (2016) demonstrated that humans can even learn to control particles from the realm of quantum physics. Not only could humans learn to control the movement of quantum atoms in the task by Sørensen et al., but more importantly, the amount ways that humans controlled these quantum atoms outperformed solutions generated previously by extensive computer simulations (for details, see Weitenberg, Kuhr, Mølmer, & Sherson, 2011).

This work exemplified the human capacities to learn to solve a complex and novel problem, providing solutions that significantly advanced a complex quantum physics’ challenge. The trick in this work was to use learning and intuition of laypeople by providing them with an game with a visual interface through which humans could easily manipulate and explore certain properties of quantum mechanics without engaging in complex mathematics. Using this method, human learners discovered sets of solutions to the underlying high-dimensional physical problem, which could then be further advanced by state of the art computer algorithms. What was crucial in this work was that humans were able, by trial and error, to detect promising sets of solutions to the underlying problem which had not been found before through extensive computer simulations pertaining to the same mathematical problem.

What we can learn from these findings is that the human mind can adapt to complex computational environments even in a domain, namely quantum mechanics, where people most likely have zero prior experience. Usually, people are not exposed to the laws of quantum physics. This stands in contrast to tasks like face recognition or image classification, in which humans can plausibly accumulate extensive experience throughout a lifetime or through evolutionary time, that may provide the cognitive system with heuristics to solve these task. By contrast, the dynamics of quantum mechanics new to laypeople, and people can only employ analogies from everyday life to solve problems in the quantum domain, or solve mathematical equations too complex to calculate on the fly. Despite that fact that quantum mechanics is new to laypeople, we have seen that human cognitive system can learn to deal with such complex dynamics if they are presented in a visually meaningful way. Given that humans lack prior strategies to control objects in the quantum world, the question arises by which cognitive mechanisms humans can get to successful solutions for problems from the quantum physics domain.

One hypothesis that has been brought forward by several authors is that the human mind reduces the complexity of the world by selectively ignoring information (REF GG) or by simplifying the statistical structure of the task (REF classification). Another factor discussed in cognitive science is that training with tasks of higher complexity facilitates solving entirely novel tasks in different domains (aking to a domain-general ability to tackle highly complex problems). The present work investigates these cognitive correlates of learning in the novel domain of quantum mechanics.

# **Task Design**

To investigate human learning we used a high dimensional control problem from quantum mechanics (adapted from Sørensen et al., 2016). The task, which is described in the next paragraph, is both complex and novel. The task complexity stems from the high dimensionality of the solution space. XXX details about the dimensionality? Further, people can be assumed to lack direct exposure to the dynamics of quantum mechanics in every-day life.

The task was a computerized motion learning task (see Figure 1A and 1B). It consisted of dragging a wave from the left side of the screen around two obstacles into a target area on the right side of the screen (in yellow in Figure 1). Crucially, due to the drag, the wave becomes instable (illustrated in Figure 1A), and although the wave looks like a liquid the destabilization is not predictable by everyday experience but follows the dynamics of wave functions in quantum mechanics. The goal is to place the wave as stable as possible into the target area. The task involved dragging the wave function around two obstacles. The time limit was 10 seconds. If during the drag, any part of the sloshing wave hits an obstacle, the wave disappeared and they had to start dragging from the starting positon again. Participants had 100 attempts to solve the task. The task was counted as solved if they had successfully dragged the wave function into the target area three times in a row. Prior to the task, participants were familiarized with the game in three training levels, shown in Figure 1A, which they also had to complete three times in a row. The first training level involved just moving the wave function to familiarize participants with the setup. The second training level introduced one obstacle that was easy to surpass. The third training level included one harder to surpass obstacle.

The last level 4 included two obstacles. In the task design, we increased the task difficulty by setting a time limit of 10 seconds, which was selected such that a simple Artificial Intelligence using a hill-climbing optimization technique was unable to solve the task.

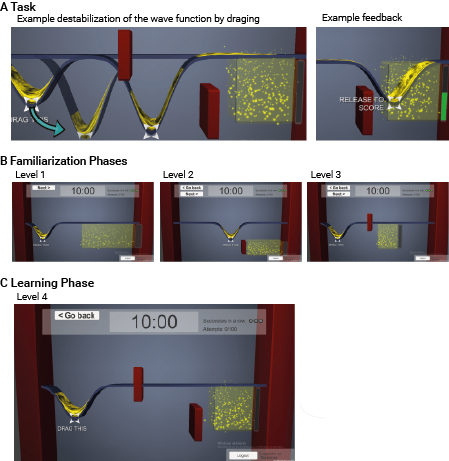


Figure 1. Experimental task. In each level, participants used the computer mouse to drag the yellow wave from the left side of the screen into the target area on the right side of the screen. Hitting one of the red obstacles in the levels 2, 3, and 4 would reset the task to zero. The Example feedback screen shows the timeout feedback.

## **Empirical Investigation**

Participants. 126 people were invited to play all four levels of the experimental game. The final sample consisted of *N*=97 people completing the experiment (29 aborted the experiment before reaching level 4 or experienced technical issues). Our final sample consisted of 51 females and 46 males; mean age 25 years (SD 6.1, range 18 to 74 years). We recruited via the COBE Lab at Aarhus University, Denmark; data were gathered from May to June 2017, the experiment was conducted according to the ethical and data protection guidelines there.

Material and Procedure. The task consisted of a computerized learning problem called QuantumMinds and psychometric measures. Initially, participants filled out the English versions of the need for cognition scale (REF) and the short NEO-PIR scale (REF) (order randomized across participants), followed by a video explaining the subsequent game, and five comprehension check questions about the game layout. Participants failing the comprehension check could re-watch the video as often as they needed until passing the checks. Afterwards, participants played the game. This game involves dragging a simulated quantum atom, which was displayed as wave, by moving the computer mouse across the computer screen into a designated target area in maximally 10 seconds. Moving the wave de-stabilizes the wave according to the laws of quantum physics, and the goal was to keep the wave as stable as possible. Participants received feedback about the final stability of their wave in the target area, which was scored between 0.00 (very instable) to 1.00 (very stable). This score is referred to *fidelity* of a path, henceforth. Participants’ task was to find, by trial and error, the best path along which dragging the wave resulted in a high fidelity, i.e. a stable wave. Participants completed three training levels of increasing difficulty, and the final critical level 4. They could play maximally 100 attempts (= trials) per level. A player continued to the next level when three consecutive paths of this player reached a fidelity above 0.50, or when exceeding the 100 attempts limit. After the game, participants filled in a flow scale (REF), demographics, and gaming experience questions.

## Results

We first describe human performance in the QuantumMinds learning task. For the generalized linear modeling the lme4 package (Bates and Maechler, 2010) and the afex package (REF) were used in the statistical environment R (version 3.4.0, R Development Core Team, 2010). Scripts and data are available as a supplement and at http://…..

Descriptive performance in the four levels of the game. The difficulty of the task increased throughout the levels of the game. To complete levels 1, 2, 3, and 4, participants used on average 12, 6, 62, and 83 attempts, respectively, of the available 100 attempts per level. The quality of their solutions, measured as the fidelity after dropping the wave by releasing the cursor, equaled on average 0.36, 0.54, 0.18, 0.09 in levels 1, 2, 3, and 4 (respectively) with a maximum fidelity equal to 1.00. Table 1 further describes frequency and quality of play across levels of the game. Participants needed more attempts to achieve the criterion in higher levels, in the order L1 ~ L2 < L3 < L4 for levels 1, 2, 3, and 4 respectively (One-way ANOVA of attempts needed by level, F(1, 3) = 234.87, p < .001, partial = .65, with 90% *CI [.60, .68]*), Table A1 in the appendix shows the full results.

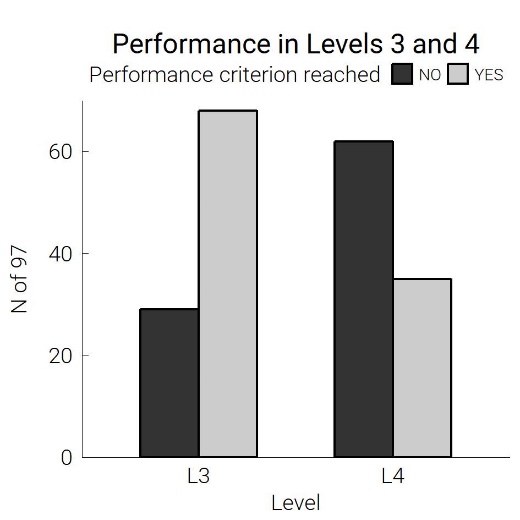
Table 1

*Fidelity and number of plays in the game*

| Level |  | Attempts | | | | |  | Fidelity | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Mean | Median | *SD* | Min | Max |  | Mean | Median | *SD* | Min | Max |
| L1 |  | 12 | 6 | 17 | 3 | 100 |  | 0.359 | 0.137 | 0.389 | 0 | 0.989 |
| L2 |  | 6 | 3 | 10 | 3 | 100 |  | 0.535 | 0.619 | 0.290 | 0 | 0.947 |
| L3 |  | 62 | 62 | 34 | 3 | 100 |  | 0.177 | 0.033 | 0.262 | 0 | 0.992 |
| L4 |  | 83 | 100 | 28 | 10 | 109 |  | 0.088 | 0.001 | 0.194 | 0 | 0.885 |

*Note:* N = 97, fidelity denotes quality of a play and ranges from 0 to 1, attempts denote the number of attempts before starting the next level.

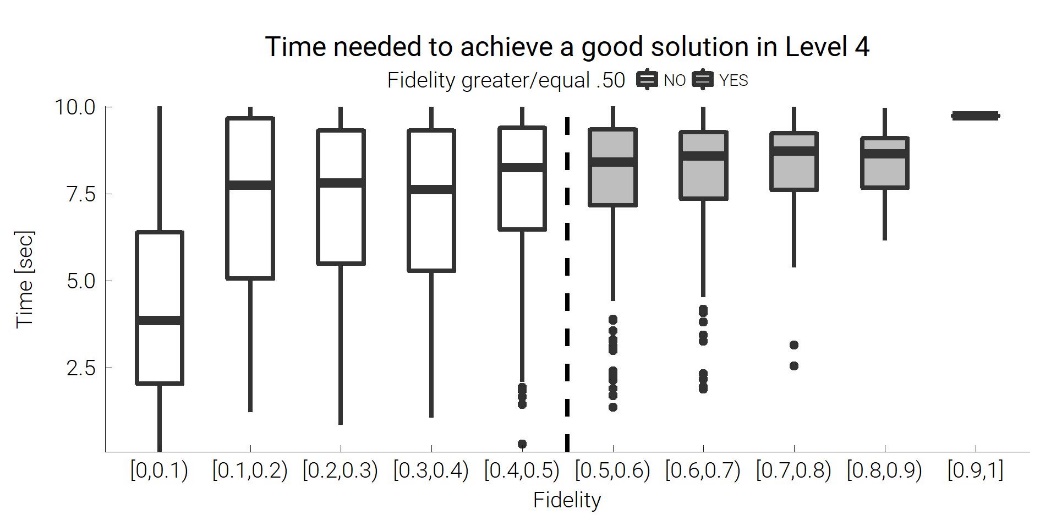
Performance in the critical 4th level 4 of the game: Humans beat a simple AI. Any AI solving this task has the advantage that the computer is not subject to subtle movement errors which humans make when dragging a mouse along a keyboars. Further, a computer can replicate its past solution exactly with minor perturbations, while repeating the exact same movement of the mouse across the screen is extremely unlikely for human hand movements. However, human players could solve the quantum game challenge in a time span that a simple AI could not: The challenge consisted in achieving a stability of 50 % of the wave after circumnavigating two obstacles. Although generally speaking, level 4, was more difficult than the last training level 3—while in level 3 a total of 15 % of participants’ attempts reached a fidelity greater/equal than .50, only 4 % of all attempts reached this fidelity level 4 (910 of 6’033 attempts in level 3, 651 of 15’826 attempts in level 4). Importantly, participants could solve the challenge despite the fact that human hand movements are more jagged than smooth computer-generated curves and that exact replication of a previous movement path is sheer impossible. Moreover, participants could replicate their successful solutions multiple times in a row. Of the 97 human players, who were instructed to replicate successful movements three consecutive times, almost half achieved this learning criterion, as shown in Figure 2. The figure further shows that in level 3 total of 68 participants (70 % of 97) achieved the performance criterion, compared to 35 participants achieving criterion performance in level 4 (36 % of 97). Modeling participants’ probability to achieve the learning criterion in level 3 compared to level 4 revealed that it was significantly less likely to achieve the criterion in level 4 compared to level 3, GLM of level 3 and 4 on success probability, change in success probability *(level 3 vs. level 4) = -1.751, SE = 0.420, p < .001* with parametric bootstrap, using a binomial link, participant as random effect, estimated by ML, and Tuckey Contrasts.



*Figure 2*. *Difficulty of levels 3 and 4 compared.* Number of participants reaching the performance criterion, defined as three consecutive fidelities greater/equal .50, in both levels.

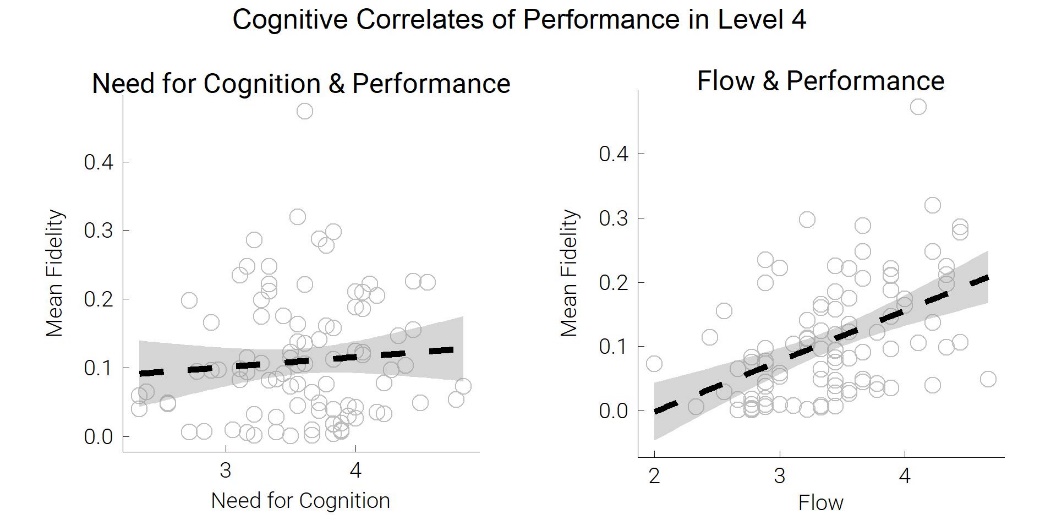
The fact that level 4 was more difficulty was expected, since we designed the critical level 4 to be novel and complex. What we were interested in is whether humans could solve the task at all, which was the case.

Given that one of the difficulties in moving the simulated quantum atom consists in the fact that the wave function becomes instable when it is moved very rapidly, it is particularly interesting how quickly participants could drop the wave off while keeping is rather stable after circumnavigating both obstacles. While keeping the wave function stable is rather easy if it is dragged on a straight line, the fact that participants had to circumnavigate two obstacles rendered this tactic impossible. Further, while taking a lot of time also eases keeping the wave function stable, the time limit of 10 seconds also rendered this strategy impossible. The fastest successful attempt that participants’ achieved (fidelity > 0.50) in level 4 took only 4.41 seconds of maximally 10 seconds, and the average duration of successful attempts in level 4 was 9.71 seconds (Med = 8.71, range = 4.41 to 10 seconds). By contrast, the unsuccessful attempts in level 4 were considerably faster than successful ones (MW = 4.90, Med = 4.50, SD = 2.96, range = 0.08 to 10.02 seconds). Figure 3 shows that human participants could outperform the computer in terms of moving the wave function around the obstacles while keeping it stable in less than 10 seconds. It also shows that more time was needed to keep the wave more stable (higher fidelity on the x axis). Importantly, players were able to achieve a very good stability of the wave with fidelities above .90 in less than 10 seconds.



*Figure 3*. *Speed of human participants to achieve a high-quality solution in level 4.* Seconds in which participants could bring the wave function into the target area depicted in Figure 1A keeping the wave stable. *Fidelity* denotes a wave function’s stability between 0.00 and 1.00. A simple AI computer algorithm needed at least 10 seconds for a fidelity greater/equal to .50 in level 4.

Cognitive Abilities vs. Motivational Factors as Drivers behind Human Performance. The excellent performance of the participants could be caused by higher concentration and more experience in solving demanding cognitive tasks. Alternatively, motivational factors could lead to higher emersion in the game. The latter hypothesis was formed since in pre-tests of the game, multiple players reported to have, at some point, stopped to think about the reaction of the wave and to have “followed their gut feelings” in order to succeed. To test if cognitive abilities or motivational factors were related to a better performance in the game the mean quality of the solution of a player was regressed on a measure of cotnitive abilities (need for cognition) and a measure of motivation (flow), while controlling for the experience that players had with computer gaming. The results revealed that emersion into the game increased performance ( = +0.07, standardized = +.44, *p <* .001), and gaming experience increased performance ( = +0.01, standardized = +.20, *p* = .028), linear Regression, = .273, but need for cognition had no significant influence on the mean quality of participants’ stabilities of the wave. Figure 4 illustrates these results, showing the marginal relationship in terms of the correlation between flow and need for cognition with the performance in level 4 of the game.



*Figure 4. Fun and emersion in the game (flow) relate to performance in level 4.* The figure illustrates the marginal cognitive correlates of participants’ good performance (without controlling for other variables, for the regression coefficients, see text). The *Mean Fidelity* denotes a performance measure from 0 to 1 with 0.00 denoting a participant never transporting the wave function into the target area stable, and 1.00 denoting a participant transporting the wave function perfectly stable into the target area in every attempt.

## Exploration Exploitation as a Mechanism that leads to Success in Level 4 of the Game

Results still unclear – no correlation between the variability in the paths that players produce at the beginning with their success in level 4 or at the end of level 4. Trying various attempts to smooth the data to get rid of artifacts.

# Discussion

The previous work investigated the mechanisms of how humans learn to solve a completely novel, complex task with a particularly high dimensional response surface, namely moving a simulated quantum atom around two obstacles as quickly as possible. The atom had to be moved within maximally 10 seconds, which was the shortest duration in which a simple Artificial Intelligence algorithm using a hill-climbing optimization routine could achieve a relatively good solution. Our results show that humans,

# Appendix

## Appendix A: Further Results

Table A1 shows the results of regressing level on the number of attempts that participants in the experiment used in each of the training levels (L1 – L3) and the final level L4.

Table A1

*Fixed-Effects ANOVA results using Number of Attempts needed as the criterion*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Predictor** | **Sum Squares** | ***df*** | **Mean**  **Square** | ***F*** | ***p*** | **partial η2** | **partial η2**  **90% CI [LL, UL]** |
| (Intercept) | 14160.66 | 1 | 14160.66 | 24.25 | .000 |  |  |
| Level | 411463.53 | 3 | 137154.51 | 234.87 | .000 | .65 | [.60, .68] |
| Error | 224243.48 | 384 | 583.97 |  |  |  |  |

*Note.* Dependent variable = Number of attempts participants used in each level. LL and UL represent the lower-limit and upper-limit of the partial η2 confidence interval, respectively.

## Appendix B Preprocessing the mouse movement trajectories

To minimize measurement errors in the mouse movement data during the gameplay, a LOESS smoothing algorithm was applied to the x-y-time series, using a bandwidth of XXX to minimize noise in the mouse movement data (REF to R package). For trajectories that are densly sampled, which is the case in our data which was collected in intervals of X, smoothing is not expected to distort the data (O’Haver, 2016).