# Microstructure Design Using a Human Computation Game

Christopher W. Adair[a,*], Hayley Evans[a], Emily Beatty[a], Derek L.Hansen[a], Seth Holladay[a], Oliver K.Johnson[a]

[a]*Department of Mechanical Engineering, Brigham Young University, Provo, UT 84602, USA*

**Abstract**

The design of microstructures that are optimized for a given engineering applications requires exploration of a rough and high-dimensional configuration space. Gradient-based algorithms are efficient, but suffer from a propensity to get stuck in local minima. Global-optimization algorithms are better at finding global minima, but are generally slow to converge. We developed and tested a Human Computation Game (HCG) for microstructure design where players interactively manipulate the microstructure to optimize an effective macroscopic material property. We investigate the impact of various game mechanics on solution quality and efficiency, and compare the HCG player solutions to those of a traditional global optimization algorithm—Simulated Annealing (SA). We show that organizing players into Synchronous teams performed better on more complex problems on average than players working Asynchronously or Solo. We also show that in the best cases, players can find microstructures that outperform those obtained by SA by up to 25% using the same number of computations, or achieve the same performance using up to 307 times fewer computational steps. By studying the optimization strategies employed by HCG players, we anticipate that improved optimization algorithms for microstructure design (and other configurational optimization problems) can be developed.

*Keywords:* Grain Boundary, Grain Boundary Networks, Human-Computer Interaction

## 1. Introduction

Materials design studies have utilized grain boundary engineering (GBE) to great success in recent years. Examples of this include increasing the thermoelectric figure of merit (ZT) of $Bi_2S_3$ by a factor of five [1], avoiding cracking in nickel-based superalloy additive manufacturing [2], and increased cycling stability of nickel-based cathodes in batteries even at high voltages [3]. These successes show the promise in GBE and suggest that many more such successes remain to be discovered [4–7].

Simulation has increasingly been considered for GBE as high throughput and big data tools have been developed [8]. In particular, structure-property models for GBs have been improving to encapsulate all 5 crystallographic degrees of freedom [9–11], and techniques have been developed to analyze networks of boundaries and their effects on macroscopic properties [12].These successes open the way for improved mesoscopic materials design through simulation.

However, significant computational challenges still remain. The number of parameters needed to fully define the network of GBs in a polycrystalline microstructure is large, making the design problem very high-dimensional. Moverover, the multitude of local minima and maxima make the optimization landscape rough [13], which makes it difficult to employ efficient gradient-based optimization algorithms.

Stochastic global methods, such as simulated annealing (SA), are useful for addressing rough opti-

*Corresponding author.
*Email address:* cwa367@byu.edu (C.W. Adair).

mization landscapes, but they are inherently slow to converge, requiring increased computational resources. This slow convergence is exacerbated by the high-dimensionality of the design problem [9, 11, 14, 15].

One pathway for solving the problem of high-dimensionality is to construct dimensionality reducing heuristics to minimize the number of degrees of freedom [14, 16]. In the context of high-dimensional configurational optimization problems that appear in other fields (e.g. biology, quantum computing), studies have been conducted using human input and intuition as sources of these heuristics [17–19]. Notable examples of these studies include the citizen science video games Foldit [20], in which players solved the folding of a specific protein better than contemporary solutions, and Quantum Moves [21], in which players solved quantum computing tunneling problems resulting in a low level heuristic that optical tweezer movement optimizations could leverage.

To obtain these heuristics, the studies framed the underlying problem of interest as a video game. A major benefit of this approach is that it provides large data sets of user output. However, obtaining quality data and efficient use of resources (human, computational, and developmental) is challenging and requires careful construction of the video game platform [17, 22–24].

In this work, we solve a microstructure design problem using a human computation game (HCG). To identify effective game mechanics for this application and to utilize human resources most effectively, we test the effects of several multiplayer collaboration modalities, and the availability of Computational Assistance (defined later), on the quality and efficiency of the resulting microstructure design solution. We find that, on average, giving players Computational Assistance results in the greatest increase in *Solution Quality*, and that players working together synchronously generate better solutions on more complex microstructures. We also find that, compared to SA, the HCG players' microstructure design solutions are up to 25.04% higher quality (for the same number of computational steps) and require up to 307 times fewer computational steps (for the same *Solution Quality*).

## 2. Background

The use of video games in collecting scientific data has increased in recent years, studied under the title of Human Computation Games (HCGs), Games with a Purpose, or Serious Games [17, 19, 25–33].

This is different from *crowdsourcing* applications such as Amazon's Mechanical Turk (MTurk), where groups of MTurk users complete small, compartmentalized tasks for small monetary compensation [29, 34–36]. Crowdsourcing tasks can include data entry and classification tasks, as well as surveys and focus groups. Each of these tasks do not need to be a complete task by themselves, and multiple contributions can be made necessary to complete the underlying problem. One problem with crowdsourcing is the motivational intent of the human inputting the data. Studies have found that MTurk users will often create solutions or answers that minimize the human effort required to generate them, and that increasing monetary rewards does not guarantee quailty solutions [18, 35]. This makes crowdsourcing a poor fit for optimization problems as the users will not willingly spend extra effort to generate high quality answers [29]. The ability to generate this intrinsic motivation is one of the reported advantages of HCG's, as players willingly spent more time in a HCG formatted problem than a purely data driven citizen science problem [23].

Much of the current body of HCG research focuses on *classification* problems such as computer vision, label sets, and visual data classification [17, 37, 38]. CAPTCHA is an example of classification research, where user input is used to both classify objects in an image and validate user input against a gold standard [39].

Many early HCGs had similar objectives, where users were given classification tasks under conditions of interest. A study by Gundry and Deterding highlights that specific design decisions such as interfaces, error checking, and aesthetic elements affect not only game outcomes such as player engagement and retention, but also scientific outcomes such as classification error percentages and task completion time [40]. Other studies reinforce this

conclusion that game mechanic design choices impact the quality of data generated by users [35, 41–44].

Design choices during the creation of HCGs have been shown to change the quality and efficiency of player output [22–24, 41]. An example of this is a study by Prestopnik and Tang, which investigated the effect of narrative elements on user engagement and output quality [23]. They found that narrative elements could increase player engagement, leading to longer player retention, and player competence, which positively affected their perception of difficulty. Gaston and Cooper studied the effects of showing visual feedback on user performance [45]. They found that players given visual feedback used fewer actions and replayed levels more often to achieve better results. Multiple literature reviews state that successful encapsulation of a research task within a game is highly dependent on the research task itself, but agree that tasks involving randomness, autonomy, and engaged problem solving can be a good match for gamification [24, 29, 40, 46].

Building on this foundation of HCG design studies, multiple studies have successfully utilized HCGs to solve high dimensional configurational optimization problems. One notable example is the protein folding game Foldit [20], mentioned earlier. The game used user inputs to manipulate protein structure models to minimize the Rosetta energy function, which in theory represents a protein's structural morphology [47]. Game players successfully constructed a folded M-PMV protein structure which had previously eluded computation due to the high dimensionality [20, 47].

In cooperative gaming, studies have employed three strategies for leveraging players as separate computing processes: Solo play, Asynchronous play, and Synchronous play. This may be compared to serial vs parallel computation in traditional computation, with a single person acting as a processor core. Solo play uses large player numbers as redundant computation, taking the best solution from the population [19, 26, 35, 48]. Quantum Moves uses this strategy, as players depend on their own skills to compete for the best scores, unable to see other players' strategies beyond the quality metric result [21]. Asynchronous play allows players to see and build on other players strategies sequentially. This collaboration modality is employed in games such as in Foldit and EyeWire [46], where groups work together sequentially to create the best quality solutions and check each other's strategies [20, 33, 47]. Synchronous play is found in most modern multiplayer games, such as Little Big Planet [49] and Minecraft [50], where players simultaneously work together in the same space to complete goals. It is currently unknown which, if any, of these strategies have the most positive effect on optimization HCG results, and if these strategies interact with previously studied game mechanics.

Criticisms of using HCGs for data generation include that while certain scientific problems are a good match for gameplay, the gameplay may not contribute useful information to the scientific problem [22, 23, 29]. Of specific note, player results from Quantum Moves did not outperform a stochastic gradient ascent algorithm in a study by Sels [21]. They concluded that if a game is to be used to solve optimization problems, it's mechanical construction should produce data that enhances or improves optimization algorithms or is generalizable to similar problems.

Taking these successes and criticisms into consideration, we have developed an HCG, *Operation: Forge the Deep*, in which players manipulate polycrystalline microstructures to optimize GB sensitive properties. In this work, we use this HCG platform to answer the following questions:

- Can the quality and/or efficiency of microstructure design solutions be improved through the use of HCGs?
- Are better and/or more efficient microstructure design solutions obtained when Solo, Asynchronous, or Synchronous collaboration modalities are employed?

## 3. Methods

### *3.1. Microstructure Design Problem*

For this study, the physical phenomenon of interest is intergranular diffusion. The objective of

the microstructure design problem is to manipulate the microstructure of a 3D polycrystal to maximize its effective diffusivity in a type-C kinetic regime—where diffusion is restricted to the grain boundary network (GBN) [51].

For each design problem (microstructure), the geometry is fixed and the design variables are the crystallographic orientations of each constituent grain. The orientation of a given grain, $A$, is defined using the quaternion parameterization,

$$q_A = [\cos(\omega_A/2), \sin(\omega_A/2)\boldsymbol{r}_A] \tag{1}$$

where $\boldsymbol{r}_A$ and $\omega_A$ are the rotation axis and angle, respectively.

The misorientation between two grains is defined as

$$q_{AB} = q_A^{-1} q_B = [\cos(\omega_{AB}/2), \sin(\omega_{AB}/2)\boldsymbol{r}_{AB}] \tag{2}$$

and the disorientation, $\hat{q}_{AB}$, is the misorientation having the smallest rotation angle among all symmetrically equivalent misorientations, and possessing a rotation axis lying in the standard stereographic triangle [52, 53].

When the orientation of a particular grain is changed, the lattice disorientations of the GBs surrounding it are modified and consequently their properties—diffusivities in this case—also change, subject to the chosen structure-property model. Changes in the diffusivities of the constituent GBs finally result in a change of the effective diffusivity of the polycrystal. Thus, as the grain orientations are manipulated the effective diffusivity ($D_{eff}$) of the polycrystal is modified. In the absence of a published 5D GB diffusivity model, a linear toy model was used to facilitate testing of the HCG method. We define this toy model as a function of GB disorientation angle and GB plane normal:

$$D(\hat{\omega}_{AB}, \boldsymbol{n}_{AB}) = \alpha(\hat{\omega}_{AB} + |n_{ABx}| + |n_{ABy}| + |n_{ABz}| + \beta) \tag{3}$$

where $\boldsymbol{n}_{AB}$ is the unit plane normal in the lab frame and $\hat{\omega}_{AB}$ is the disorientation angle. The constants $\alpha$ and $\beta$ are arbitrary scaling factors chosen to have values of $10^7$ and 1, respectively, so that the *Total Score* (See Section 3.3.3) had values greater than 1 and maximum values less than 10.

### *3.2. Model Definition and Simulation Setup*

Non-periodic synthetic 3D polycrystals, respectively containing 10 or 20 Grains, were generated using the Neper polycrystal generation/meshing software [54]. Combined Dirichlet and Neumann boundary conditions were applied, as illustrated in Fig. 1a, so that the macroscopic diffusion flux occurred in the positive $x$ direction.

The effective diffusivity of each polycrystal was calculated using the finite volume method (FVM) employed in Johnson et al. [12]. A surface mesh of the GBN was constructed using the frontal meshing algorithm in Neper (which promotes low-aspect ratio elements) [54], as shown in Fig. 1c. The geometry and position dependent properties of the GBN mesh are encoded in the GBN Laplacian matrix [12, 55, 56]:

$$\mathcal{L}_{ij} = \begin{cases} \sum_{i \sim m} \dfrac{D_{im} A_{im}}{L_{im}} & \text{if } i = j \\ -\dfrac{D_{ij} A_{ij}}{L_{ij}} & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\mathcal{L}_{ij}$ is the $ij$-th element of the GBN Laplacian matrix, $D_{ij}$ is the GB diffusivity from Eq. 3 assigned to the edge connecting mesh vertices $i$ and $j$ (see Fig. 1c), $L_{ij}$ is the edge length, and $A_{ij} = W_{GB} c_{ij}$ is the effective cross-sectional area of the $ij$-th edge (through which the flux between nodes $i$ and $j$ flows), with $W_{GB}$ a constant GB thickness and $c_{ij}$ being the distance between the centroids of the incident triangular elements, as shown in Fig. 1d. For simplicity, we neglect diffusion along the triple junctions, and we assume steady-state conditions. The effective diffusivity of the microstructure was calculated using the spectral formulation of the FVM according to [12]:

$$D_{eff} = \frac{L}{A} \left( \sum_{k \geq 1} \lambda_k^{-1} [u_k(a) - u_k(b)]^2 \right)^{-1} \tag{5}$$

where $D_{eff}$ is the effective diffusivity of the polycrystal, $L$ is the length of the microstructure in the $x$-direction , $A$ is the cross-sectional area of the microstructure orthogonal to the $x$-direction. $\lambda_k$ and $u_k$ are the eigenvalues and eigenvectors of the

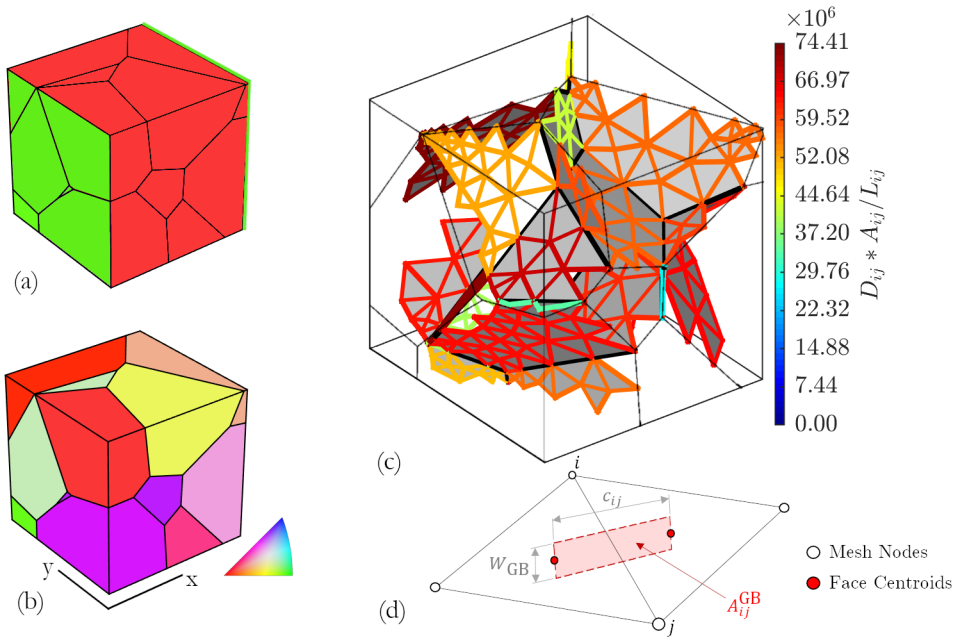



Figure 1: (a) Boundary conditions for diffusion through the simulation cell. The green planes employ the Dirichlett condition with fixed concentrations, while all other planes (red) have Neumann (zero-flux) boundary conditions. (b) Microstructure with grain orientations assigned. (c) Meshing of the internal GBs, with the resulting property colored on the edges (d) Representation of an edge element with the dual of the mesh used for FVM calculations.

Laplacian matrix respectively, and $a$ and $b$ are the combined Dirichlet condition diffusion source and sink nodes, respectively (see [12]).

### *3.3. Solution Methods*

The solution of the design problem consists of assigning crystallographic orientations (Fig. 1b) to each of the constituent grains such that the effective diffusivity is maximized. This constitutes a high-dimensional configurational optimization problem with $3N_{grains}$ degrees of freedom, where $N_{grains}$ is the number of grains in the polycrystal, and each requires 3 parameters to specify the 3D rotation defining its crystallographic orientation.

#### *3.3.1. Simulated Annealing*

We employed two different approaches to solve this microstructure design problem. The first was a simulated annealing (SA) algorithm. This is a common stochastic global optimization algorithm employed for problems with large configurational spaces such as GBNs [13, 57] and was used as a gold standard to compare against the HCG results.

In SA, at each Monte Carlo (MC) step, a random grain is selected and assigned a new orientation. The new orientation was selected in one of two ways (to facilitate comparison with the HCG): (i) uniformly at random, or (ii) via a local gradient ascent that maximized the diffusivities of the incident GBs (i.e. without consideration of the global effective diffusivity).

The new grain orientation assigned at each step was accepted if either the evaluated objective function ($D_{eff}$) improved, or if a randomly-generated number fell within the current acceptance interval defined by the annealing schedule [13]. We employed a convergence criterion of 1000 rejected steps.

#### *3.3.2. Human Computation Game*

The second solution strategy was the HCG that is the focus of this work. The goal of solving this microstructure design problem via an HCG is to leverage human 3D spatial reasoning [19, 20] and intuited dimensionality reducing heuristics [15]. However, players are anticipated to be non-experts with no background in crystallography,

microstructures, or grain boundaries. Therefore, the microstructure optimization problem needed to be abstracted and represented in an understandable manner while retaining the required crystallographic inputs.

To facilitate the interpretation of the problem by non-experts, we simplified the visualization of the full 3D microstructure (Fig. 2a) to a minimalistic representation (Fig. 2c). The orientation of each grain was represented by a cube (reflecting cubic crystal symmetry) positioned at the grain center. GBs were implicitly represented by a "connection" (rods shown in Fig. 2c) between adjacent grain centers. Admittedly, to an expert this may appear confusing in that it might suggest transgranular transport; however, to non-experts this visually communicated the dependence of the boundary's property on the neighboring grains' orientations.

To visually communicate the magnitude of the diffusivity of a GB between two grains (and how users' moves affect it), the length, diameter, and color of these connections were set to scale proportionally to the respective GB diffusivity. Note that when a user rotates a grain's orientation cube, all of the incident connections are updated simultaneously because all of its surrounding GB disorientations, and hence GB diffusivities change. A full view of the in-game user interface can be seen in Fig. 3.

Compared to a direct literal representation of the polycrystal (Fig. 2b), this simplified structure (Fig. 2c and Fig. 3) achieved greater visual clarity and learnability for players in our early testing, which is why it was chosen for our final implementation.

To play the game, players select one of the available cubes, which is then highlighted yellow for identification (see Fig. 3). Players then rotate their chosen cube by clicking and dragging the mouse. As described earlier, as the grain's orientation changes, the disorientations, and consequently diffusivities, of its GBs are updated. The effective diffusivity of the entire polycrystal (which is the underlying objective function) is also updated and is displayed to the user as the *Total Score* (see Fig. 3), which they attempt to maximize by iteratively repeating this procedure (selecting a cube and then rotating it, selecting another cube and rotating it, etc.).

In addition to manual rotation, players are also presented with an *Optimize Current Grain* button (labeled *Auto* in Fig. 3). If a player clicks this button, a gradient ascent algorithm is applied to the currently selected grain, which attempts to rotate it into the orientation that maximizes the diffusivities of the incident grain boundaries (i.e. it performs a local optimization step on the currently selected grain). Importantly, this is done without regard for the *Total Score*, so the resulting orientation may be locally optimal, but globally sub-optimal, since the interaction between local cube orientation and quality of the global configuration is complex, and changes made to a single grain will affect multiple grain boundaries (all those that are incident to it). A similar mechanic of optional Computational Assistance was employed in the game Foldit, where the computer agent is called, on-demand, to find local solutions [20, 47].

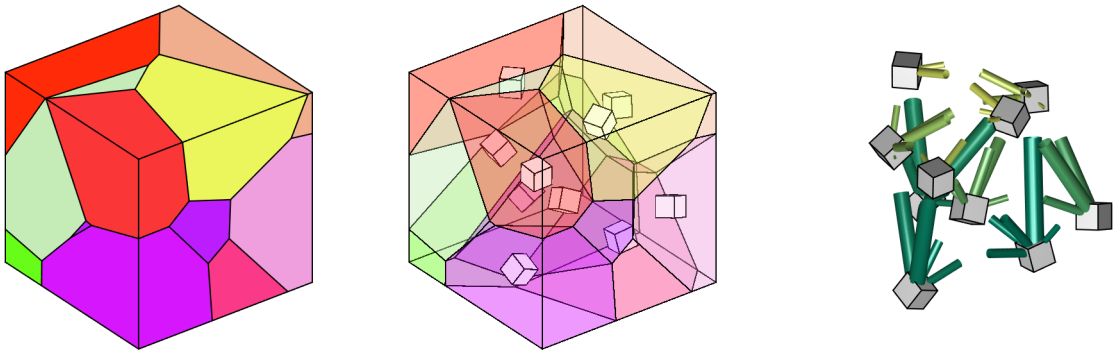
Prior to beginning a game session, players completed a 5 minute in-game tutorial, where they learned how to select and rotate the cubes, how to use the *Optimize Current Grain button*, and became familiar with the visual and numerical feedback mechanisms.

#### *3.3.3. Quantifying Human Performance*

We employed two metrics to quantify the performance of the HCG (and the SA algorithm) in solving the microstructure design problem: *Solution Quality*, and *Computational Efficiency*. The *Solution Quality* is simply the value of the objective function, $D_{eff}$, (displayed to the user as the *Total Score*, as shown in Fig. 3). The *Computational Efficiency* was evaluated at the end of a game, and is defined by:

$$\eta = \frac{Total\ Score}{N} \tag{6}$$

where $\eta$ is the *Computational Efficiency* and $N$ is the number of computational steps (i.e. the number of times the objective function was evaluated). The objective function, $D_{eff}$ or *Total Score*, was evaluated only (i) when the player committed a move by lifting their mouse button after rotating a cube, or (ii) after the *Optimize Current Grain*



(a) (b) (c)

Figure 2: Representation of grain boundary networks. (a) A simulated 3D polycrystal generated with Neper Polycrystal [54]. (b) A simulation with cubes at the grain centers representing the respective grain orientations. (c) Abstracted visual representation of the microstructure in which the "connections" between cubes represent the magnitude of the property on the corresponding GB defined by those two grains.


Figure 3: Full game view of a 3D grain boundary network in-game. Players are able to maneuver through the 3D space to manipulate the orientations of grains, and receive visual feedback on how their changes affect both the local properties of the incident GBs and the effective macroscopic material property. The view shown here represents current version of the HCG. Players participating in the study were given a visually simplified yet functionally identical version.

*button* finished its action (recall that this button activates a local optimization routine, which never evaluates the global objective function). The value of $\eta$ is essentially a measure of the average improvement in the objective function per computational step—which may reduce with gratuitous or random decisions and may increase with effective strategic decisions.

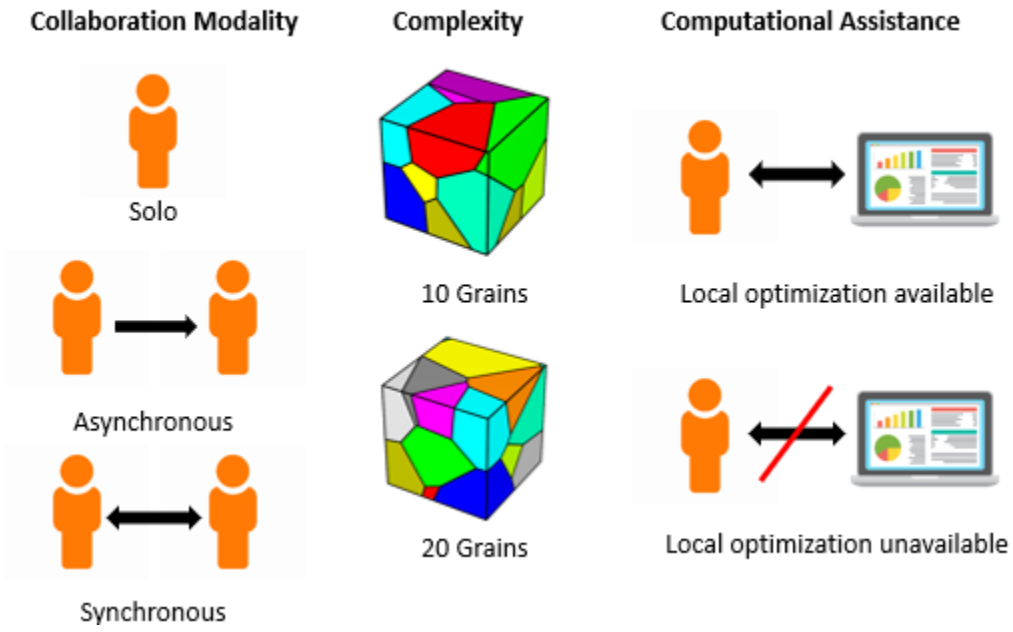



Figure 4: Visual representation of the full factorial study, including the three factors (collaboration modality, problem complexity, and Computational Assistance), with each of their respective levels enumerated.

### *3.4. Experimental Design*

To answer our two motivating questions, we fitted a Linear Mixed Model (LMM) of three categorical variables (2-3 levels each) and their interactions as shown in Fig. 4, with the dependent variable being $D_{eff}$ (when evaluating *Solution Quality*) or $\eta$ (when evaluating *Computational Efficiency*), respectively. A LMM was employed due to the presence of both fixed and random variables.

The factors included the multi-player collaboration modality (with levels corresponding to Solo, Asynchronous, or Synchronous), the complexity of the puzzle (with levels of either 10 or 20 Grains), and the availability of Computational Assistance (with levels of true or false, corresponding to whether or not the *Optimize Current Grain* button was available to the player). The coefficients of this model represent the estimated contribution of each factor level (or their interactions) on $D_{eff}$ (or $\eta$).

The multi-player collaboration modalities were respectively defined as follows. Solo play consisted of a single player manipulating the microstructure. In Synchronous collaboration, two players were connected to the same microstructure design problem hosted on a remote server and simultaneously manipulated the microstructure. The players could select and rotate any cube except the one that their partner was actively manipulating. In Asynchronous collaboration, each player began working on a puzzle, and then swapped with their partner after half of their allotted time had elapsed. The players then built on their partner's work attempting to further optimize the microstructure in the time remaining.

The complexity of each microstructure design problem was approximated by the number of grains in the microstructure, which defines the dimensionality of the configuration/design space. This represents a simple test for the scalability of human input against simulation size. For this study, players were presented with microstructures containing either 10 or 20 Grains.

We used a randomized blocking structure to organize subjects into a mixed study. Each group of two subjects in a session played every combination of complexities and collaboration modalities in random order (within-subjects). Each group was randomly placed in either a Computational Assistance available condition or a Computational Assistance not available condition (between-subjects). Two groups, one with and one without Computational

Assistance, made up a single block in the randomized blocking structure (consisting of a total of 4 players). This was done to mitigate the learning effect, block the effects of individual groups, and reduce the number of subjects needed for statistical power [58]. In every test condition, players had 8 minutes to complete a given microstructure design problem.

Players were recruited strictly on a volunteer basis, with no compensation given. Players were instructed on gameplay through an automated tutorial and proctored through a session by script. During all sessions, player input (i.e. each action taken in the game) was captured in the background for future reconstruction and analysis.

A session on average took 50 minutes to complete. There were a total of 32 participants, taken from the student body of a large university. This led to a total of 96 data points, or 9 replicates of every experimental condition. After completion of the study, participants were given an exit survey to collect qualitative feedback. Due to the class of IRB approval employed, no demographic data was collected from subjects. We do not anticipate that variations in demographics impact the results in any significant way.

After data collection was completed, the data was analyzed using the LMM. The model included a constant intercept, all three fixed effects (see Fig. 4), their interactions, and the blocking variable (which group a participant was assigned to), which was designated as a random variable.

## 4. Results

The LMM was analyzed using a reduced maximum likelihood (REML) method. Analysis of the statistical model found significant effects in the factors for both the *Solution Quality*, and the *Computational Efficiency*.

### *4.1. Solution Quality*

Table 1 shows the LMM coefficients and the statistical significance ($P$-value) of the effect of the various factors (and their interactions) on the *Solution Quality*. The intercept, considered alone, represents the baseline test condition, which consisted of Synchronous collaboration, a 20 Grain microstructure, and no Computational Assistance (no *Optimize Current Grain* button). The parenthetical identifier indicates the factor level that was changed from the baseline condition, and the corresponding coefficient represents the resulting change in *Solution Quality* (i.e. the change in $D_{eff}$). For example, under the baseline conditions the model predicts that the average HCG player score would be 7.0924 (the value of the intercept). If the Computational Assistance factor is changed to "Yes" then the model predicts an increase in the average HCG player score of +0.1095 (the value of the corresponding model coefficient.). Or if Computational Assistance is changed to "Yes" and the Collaboration Modality is changed to "Solo" there is a predicted decrease in the average score of –0.0858 (the value of the corresponding model coefficient). Note that the coefficients for the factor levels of the baseline condition are by definition zero and are therefore not included in the table.

This regression had an adjusted $R^2$ value of 0.869, implying a good fit to our data. The provided $P$-value represents the result of a $t$-test where the null-hypothesis is that the respective coefficient has a value of zero (which would imply that it has no influence on the average player score). The coefficients whose values show statistical significance are the intercept, Complexity (10 Grains), Computational Assistance (Yes), and the interaction of Computational Assistance (Yes) and Collaboration Modality (Solo). The random effect (individual groups) accounted for only 3.09% of the total variance in the model.

Using Tukey honestly significant difference (HSD) post-processing [58], we can more robustly evaluate the effects of the interactions. From the post processing, Tukey HSD found no statistical differences between any of the cases with 10 Grains, suggesting that neither collaboration modality nor availability of Computational Assistance truly affected $D_{eff}$ in the 10 Grain condition. However, HSD showed that for the 20 Grain complexity level one test condition exhibited a statistically significant performance difference from the rest. Groups engaged in Synchronous collab-

Table 1: List of coefficients and $P$-values found by the LMM model for *Solution Quality*.

| Factors on Solution Quality | Model Coefficient | $P$-value |
|---|---|---|
| Intercept | 7.0924 | $<0.0001^{**}$ |
| Collaboration Modality (Solo) | −0.0694 | 0.0941 |
| Collaboration Modality (Async) | −0.0184 | 0.6545 |
| Complexity (10 Grains) | −0.7031 | $<0.0001^{**}$ |
| Computational Assistance (Yes) | 0.1095 | $0.0003^{**}$ |
| Computational Assistance (Yes) × Collaboration Modality (Solo) | −0.0858 | $0.0395^{*}$ |
| Computational Assistance (Yes) × Collaboration Modality (Async) | 0.0018 | 0.9653 |
| Computational Assistance (Yes) × Complexity (10 Grains) | 0.0060 | 0.8356 |
| Collaboration Modality (Solo) × Complexity (10 Grains) | 0.0401 | 0.3311 |
| Collaboration Modality (Async) × Complexity (10 Grains) | 0.0151 | 0.7125 |
| Computational Assistance (Yes) × Collaboration Modality (Solo) × Complexity (10 Grains) | 0.0593 | 0.1513 |
| Computational Assistance (Yes) × Collaboration Modality (Async) × Complexity (10 Grains) | 0.0462 | 0.2622 |

(*) represents significance at the 0.05 level, and (**) represents significance at the 0.001 level.

oration, which also had Computational Assistance available, were found to perform significantly better than other test conditions for the 20 Grain problems.

In summary, there is a statistically significant effect of the problem complexity with larger problems (the 20 Grain case) resulting in higher effective diffusivity than smaller ones (the 10 Grain case). For the lower complexity problem (the 10 Grain case) there was no significant effect of collaboration modality or availability of Computational Assistance (the *Solution Quality* did not show a statistically significant difference for any of the levels of these factors). Thus, for low-complexity problems the benefits of particular collaboration modalities and even Computational Assistance are not apparent (such problems can apparently be solved just as well without them). However, for the higher complexity problem (the 20 Grain case) a statistically significant improvement in *Solution Quality* was observed when the Collaboration Modality was Synchronous and when Computational Assistance was available.

### *4.2. Computational Efficiency*

Our second measure of interest was the *Computational Efficiency* of the HCG. Table 2 shows the significant effects and interactions of the factors on the *Computational Efficiency*.

This model performed much poorer than the *Solution Quality* model with an adjusted $R^2$ value of just 0.359. The blocking factor accounted for 13.6% of the total variance in this case. Tukey's HSD post-processing found significant differences only in the Collaboration Modality. There was no significant difference between the Solo and the Asynchronous cases, but the Synchronous case was statistically less efficient than either of the other collaboration modalities. This suggests that the Synchronous collaboration modality was less efficient than Asynchronous collaboration or Solo play. Thus, while differences in the effect of different factors on *Computational Efficiency* were less conclusive than their effects on *Solution Quality*, it does appear that while Synchronous collaboration leads to better *Solution Quality* on average, it does so at some cost to *Computational Efficiency*. However, the relative magnitudes of these costs and benefits must be weighed. This, together with comparison of the results for the HCG with those of SA will be described in Section 5.2.

### *4.3. User Feedback*

In the exit survey, the players were asked to select their most and least favorite collaboration modality. The players overwhelmingly voted for the Synchronous case as their favorite (55%), with the Asynchronous case being the next favorite (24%), with 10% of respondents showing no preference. The least favorite had a much narrower spread, with Solo being the least favorite (34%), then Asynchronous (27%), with 24% of respondents showing no preference. If these responses are taken as the most and least motivational cases respectively, the responses suggest that there may be a trade-off between player motivation and *Computational Effi-*

Table 2: List of coefficients and $P$-values found by the LMM model for *Computational Efficiency*.

| Factors on Computational Efficiency | Model Coefficient | $P$-value |
|---|---|---|
| Intercept | 0.1142 | $<0.0001^{**}$ |
| Collaboration Modality (Solo) | 0.0195 | $0.0036^{**}$ |
| Collaboration Modality (Async) | 0.0116 | 0.0775 |
| Complexity (10 Grains) | −0.0122 | $0.0096^{*}$ |
| Computational Assistance (Yes) | 0.0110 | $0.0184^{*}$ |
| Computational Assistance (Yes) × Collaboration Modality (Solo) | −0.0023 | 0.7260 |
| Computational Assistance (Yes) × Collaboration Modality (Async) | 0.0081 | 0.2138 |
| Computational Assistance (Yes) × Complexity (10 Grains) | −0.0041 | 0.3702 |
| Collaboration Modality (Solo) × Complexity (10 Grains) | −0.0029 | 0.6609 |
| Collaboration Modality (Async) × Complexity (10 Grains) | −0.0037 | 0.5713 |
| Computational Assistance (Yes) × Collaboration Modality (Solo) × Complexity (10 Grains) | 0.0019 | 0.7684 |
| Computational Assistance (Yes) × Collaboration Modality (Async) × Complexity (10 Grains) | −0.0106 | 0.1049 |

(*) represents significance at the 0.05 level, and (**) represents significance at the 0.001 level.

*ciency*. Specifically, players may prefer to play Synchronously, but Asynchronous and Solo modalities appear to result in more efficient solutions on average. This trade-off is somewhat mitigated by the fact that Synchronous play showed the potential to produce higher quality solutions (even if it was less efficient), thus average *Solution Quality* and player motivation may be correlated.

## 5. Discussion

### *5.1. Effects on Solution Quality*

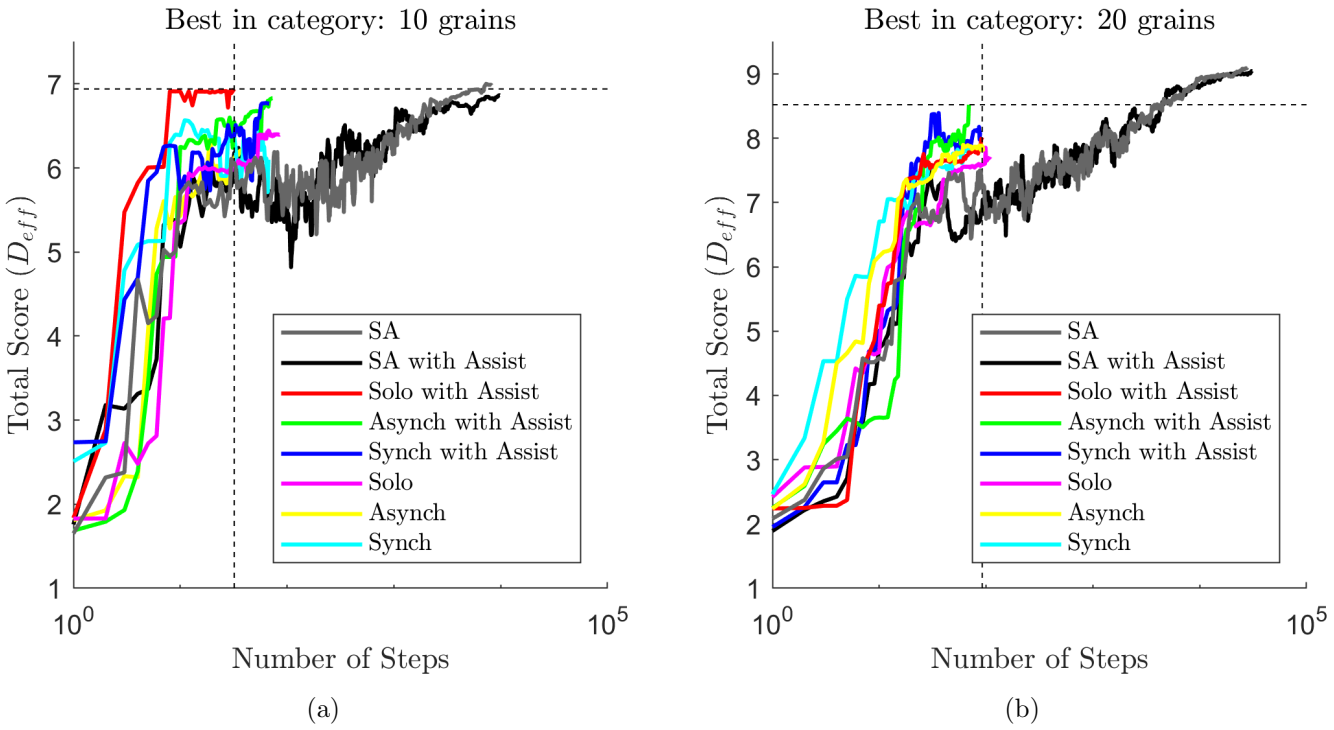
In answer to our question about whether microstructure design problems can benefit from HCGs, Table 1 and Section 4.1 show that there are sets of game mechanics that tend to produce higher quality solutions than others. In the context of our HCG, the set of game mechanics that produced the highest quality solutions on average was Synchronous collaboration with Computational Assistance applied to more complex microstructure design problems.

This optimal set of mechanics relies on the positive interaction between groups working together and the availability of Computational Assistance. Without Computational Assistance, this set of game mechanics did not produce solutions of improved quality.

The effect of the microstructure design problem complexity on the *Solution Quality* on its own, while large (0.7031), is not particularly interesting, as it simply reflects the fact that microstructures with more grains will have more GBs and therefore an increased number of pathways for intergranular diffusion, leading to a higher effective diffusivity [59]. Consequently, the more complex (20 Grain) design problems simply exhibited a higher maximum achievable *Total Score*. The players were, however, ignorant of this relationship, being non-experts, and yet were able to achieve these higher scores within the same time limit as the less complex puzzles. In other words, their performance did not decrease as the complexity increased even though the time limit remained fixed. This is additional evidence supporting Yildirim's hypothesis that time pressure can increase the quality of player inputs [60].

It is important to note that the LMM characterized the **average** *Solution Quality* (or *Computational Efficiency*) for the various test conditions. As will be discussed later, it is of even greater importance to know which conditions have the potential to produce the **best** performance (even if the corresponding average performance is suboptimal). To place the *Solution Quality* improvement into the context of microstructure optimization, we also compared the best HCG results to the best output of 5 trials of the corresponding SA algorithm (i.e. we compared the HCG with Computational Assistance to SA with Computational Assistance, or the HCG without Computational Assistance to SA without Computational Assistance).

We consider three methods for fair comparisons: equal step percentage improvement, where *Solution Quality* is evaluated when the shorter of the




(a) (b)

Figure 5: (a) Best-in-class player optimization paths, with the x axis as a logarithmic scale, compared to the best of 5 SA optimizations for 10 Grains, and (b) the same for 20 Grains. As an example, the vertical line visually defines the equal step percent quality increase, and the horizontal line indicates the equal quality percent reduction in steps for one HCG trajectory.

two optimizations end; equal quality percentage, where the ratio of steps taken to achieve the same SA quality is determined at the peak HCG quality; and percentage of maximum quality, where the maximum *Solution Quality* for both HCG and SA are considered (See Fig. 5 and Table 3). Table 3 lists the best *Solution Quality* (i.e. *Total Score* or, equivalently, $D_{eff}$) and *Computational Efficiency* achieved for each HCG experimental condition, as well as percent change relative to the best of the corresponding SA algorithms.

Fig. 5 shows the individual optimization pathways (i.e. the time history of the objective function) resulting from HCG optimization and SA. Each line reflects the exploration players did to converge to their solution. One feature of interest is the similar shape of the convergence trajectories for both SA and HCG, starting with a transient rapid increase before transitioning to a more gradual steady state (approximately linear) improvement. However, the transition to steady state for the HCG players, while occurring at a similar number of computational steps, jumped to a higher *Solution Quality* than SA in all cases. This is especially visible for the more complex (20 Grain) microstructure design problems (Fig. 5b). This higher steady state *Solution Quality* indicates that the strategies employed by the human HCG players are more effective than the SA optimization rules. This suggests the possibility of a hybrid optimization algorithm in which something like SA is used, but with the selection and modification rules (selecting grains at random and changing their orientation at random) replaced by the corresponding rules intuited by the humans. The evidence for the effectiveness of the human-intuited strategies is strengthened by the existence of HCG player trajectories that outperform SA even without Computational Assistance (i.e. the performance improvement is not due simply to the use of local gradient-based optimization).

By comparing the best performance of each test condition, we find that the best quality solutions were obtained when Solo players had assistance in the 10 Grain case, and Asynchronous players had assistance in the 20 Grain case (Fig. 5 and Table 3). The corresponding best-in-class player *Solution Quality* was 6.940 for the 10 Grain case and 8.522 for the 20 Grain case, which represents a 12.32% improvement over SA for the same number of computational steps for 10 Grains, and 25.04% improvement for 20 Grains (see the dashed vertical lines in Fig. 5 and the numerical values in Table 3). Thus, in the best quality HCG simulations, players were able to reach higher quality solutions for the same number of computational steps when compared against SA.

It is interesting to note that, consistent with the LMM predictions and the results of Section 4.1, on average the 20 Grain Synchronous collaboration condition with Computational Assistance scored 0.5 higher than the 10 Grain cases (see Table 1). However, the best solution for the 20 Grain microstructure was obtained under the other two collaboration modalities. This suggests that while, on average, Synchronous players may generate better solutions, other conditions may still generate useful, or even better, solutions. This also highlights the fact that the best performance (which is not captured in the LMM) is likely of greater interest than the average performance, especially when the conditions that maximize average performance differ from those that maximize the best performance.

### *5.2. Effects on Computational Efficiency*

In the context of *Computational Efficiency*, the most efficient solutions for a fixed *Solution Quality* were obtained under the conditions of Solo player without assistance for the 10 Grain case and Synchronous players with assistance for the 20 Grain case (see horizontal line in Fig. 5 and numerical values in Table 3).

On average, increasing complexity was correlated with increasing *Computational Efficiency* (+0.01218 Table 2). However, the highest efficiency was observed for the 10 Grain case, where the player outperformed the SA with gradient ascent results. The shape of the corresponding optimization trajectory reflects the effects of the time limit on the number of computations more than the complexity itself. In other words, with the less complex design problem, players plateaued in their progress (seen in Fig. 5), and then iterated around

Table 3: Comparison of best-in-class *Solution Quality* and *Computational Efficiency* relative to corresponding SA.

| Factor Case | Solution Quality | Solution Efficiency | Equal Step % Increase in Quality | Equal Quality % Reduction in Steps | % of Maximum SA Quality |
|---|---|---|---|---|---|
| Solo, 10, CA | 6.940 | 0.2169 | 12.32 | 816.44 | 101.07 |
| Solo, 10, NCA | 6.446 | 0.0962 | 13.62 | 30706.3* | 92.07 |
| Solo, 20, CA | 8.014 | 0.0871 | 17.46 | 778.26 | 88.60 |
| Solo, 20, NCA | 7.854 | 0.0793 | 15.65 | 367.68 | 86.37 |
| Asynch, 10, CA | 6.829 | 0.0936 | 24.42 | 877.78 | 99.47 |
| Asynch, 10, NCA | 6.361 | 0.3348 | 14.60 | 1142.11 | 90.86 |
| Asynch, 20, CA | 8.522 | 0.1235 | 25.04 | 1814.49 | 94.21 |
| Asynch, 20, NCA | 7.898 | 0.1295 | 12.53 | 829.51 | 86.85 |
| Synch, 10, CA | 6.773 | 0.1075 | 20.30 | 328.36 | 98.64 |
| Synch, 10, NCA | 6.567 | 0.5970 | 12.64 | 3700.00 | 93.81 |
| Synch, 20, CA | 8.395 | 0.2332 | 15.49 | 2769.44 | 92.80 |
| Synch, 20, NCA | 7.836 | 0.1285 | 11.64 | 596.72 | 86.17 |

(*) This is a conservative lower-bound, estimated using SA ending point because player outperformed SA given the stated constraints.

their nominal solution attempting to find small improvements to their *Total Score* with the remaining time. The increased complexity simply delays the performance plateau because there are more variables to be optimized, which in turn reduces the *Computational Efficiency*. This suggests that there may be an optimal time limit for a given complexity, and supports the results from a study on player performance by Yildirim that found evidence of an optimal time pressure for performance [60].

The most important finding is the large difference between the efficiencies of the HCG players and SA. We find that players were up to 307 times more efficient than the SA algorithm for a fixed *Solution Quality*. Thus we find that non-expert, novice players of this HCG were able to learn patterns and generate efficient strategies for solving this high-dimensional configurational microstructure design problem, reaching up to 101.07% of the gold standard's (SA) *Solution Quality*, and doing so 307 times more efficiently than SA. We therefore conclude that microstructure design problems can benefit significantly from the use of a HCG as an optimization engine.

### *5.3. Generalization*

It is important to note that the construction of an HCG represents a significant capital and time investment. Consequently, it is likely impractical to develop a custom HCG for each new microstructure design problem. In some cases, new design problems can be added to an existing HCG (like the one presented here) as new levels. For example, if the design objective changes (e.g. maximizing elastic modulus instead of GBN diffusivity, or perhaps optimizing the tradeoff between them), but the design variables are similar (grain orientations) then extending an existing HCG is straightforward. In other cases, the design problem may involve design variables that are so different that it would require impractical modification of an existing HCG, or the difference may be a thematic one such that the problem is simply out of place in the existing game.

However, we find that the optimization heuristics discovered by HCG players converge much better than SA, as seen in Fig. 5. Thus, it is still desirable to leverage the best performing algorithms developed by HCG players for a broader range of design problems than can be accommodated in a given HCG, without the cost of building a custom HCG for each problem. If the HCG player strategies could be understood and computationally recreated, then a more efficient algorithm than the traditional SA algorithm could be constructed for high-throughput GBN optimization, and other high-dimensional configurational optimization problems. The reductions to needed computational resources will allow for larger simulations and/or more complex structure-property

models. This is especially important if the underlying structure-property model or homogenization calculation is expensive to evaluate.

Thus, in future work we intend to analyze the HCG player strategies and translate the most effective strategies into novel optimization algorithms that may be generalized and applied to arbitrary configurational optimization problems, similar to the way that SA is used now. Indeed, this could lead to the most significant advancements that result from microstructure design HCGs, far beyond the results of any one HCG.

## 6. Conclusions

We have developed a human computation game (HCG) for the solution of microstructure design problems. We used this platform to investigate whether the quality and/or efficiency of microstructure design solutions can be improved through the use of HCGs compared to traditional global optimization strategies—Simulated Annealing (SA). We also investigated which game mechanics (collaboration modalities, microstructure complexities, and availability of Computational Assistance) produce the highest quality and most efficient design solutions. Our major findings include:

- Synchronous collaboration leads to increased *Solution Quality* on average, while the maximum observed *Solution Quality* was achieved by a Solo player with Computational Assistance for 10 Grains, and Asynchronous players with Computational Assistance for 20 Grains.
- In the context of a video game, non-expert human players found microstructure design solutions that performed up to 25% better than those obtained by SA, for the same number of computational steps.
- Human players were able to reach their high scores with up to 307 times fewer computational steps than SA for the same *Solution Quality*.
- Even when SA was given vastly more computational resources (i.e. considering the asymptotic performance of SA), players were still able to reach solutions whose quality was comparable to (94.21% for the best 20 Grain HCG solution) or exceeded (101.07% for the best 10 Grain HCG solution) that of SA solutions.

In summary, we observe that the optimization strategies employed by the HCG players outperform SA in both *Solution Quality* and *Computational Efficiency*. If these heuristics can be characterized and incorporated into new automated optimization algorithms it may be possible to leverage them for improved solution of a broad array of microstructure design and other configurational optimization problems.

## Acknowledgement

The material presented here is based upon work supported by the National Science Foundation under Grant No. DMR-1654700.

## Declaration of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] J. Guo, Q. Lou, Y. Qiu, Z. Y. Wang, Z. H. Ge, J. Feng, J. He, Remarkably enhanced thermoelectric properties of bi2s3 nanocomposites via modulation doping and grain boundary engineering, Applied Surface Science 520 (2020) 146341. doi:10.1016/j.apsusc.2020.146341.

[2] P. Kontis, E. Chauvet, Z. Peng, J. He, A. K. da Silva, D. Raabe, C. Tassin, J. J. Blandin, S. Abed, R. Dendievel, B. Gault, G. Martin, Atomic-scale grain boundary engineering to overcome hot-cracking in additively-manufactured superalloys, Acta Materialia 177 (2019) 209–221. doi:10.1016/j.actamat.2019.07.041.

[3] X. Cheng, J. Zheng, J. Lu, Y. Li, P. Yan, Y. Zhang, Realizing superior cycling stability of ni-rich layered cathode by combination of grain boundary engineering and surface coating, Nano Energy 62 (2019) 30–37. doi:10.1016/j.nanoen.2019.05.021.

[4] Y. Chen, C. A. Schuh, Diffusion on grain boundary networks: Percolation theory and effective medium approximations, Acta Materialia 54 (2006) 4709–4720. doi:10.1016/j.actamat.2006.06.011.

[5] V. Randle, G. Owen, Mechanisms of grain boundary engineering, Acta Materialia 54 (2006) 1777–1783. doi:10.1016/j.actamat.2005.11.046.

[6] T. Watanabe, Grain boundary engineering: Historical perspective and future prospects, Journal of Materials Science 46 (2011) 4095–4115. doi:10.1007/s10853-011-5393-z.

[7] O. K. Johnson, C. A. Schuh, Texture mediated grain boundary network design in three dimensions, Mechanics of Materials 118 (2018) 94–105. doi:10.1016/j.mechmat.2017.12.001.

[8] G. R. Schleder, A. C. Padilha, C. M. Acosta, M. Costa, A. Fazzio, From dft to machine learning: Recent approaches to materials science - a review, JPhys Materials 2 (2019). doi:10.1088/2515-7639/ab084b.

[9] V. V. Bulatov, B. W. Reed, M. Kumar, Grain boundary energy function for fcc metals, Acta Materialia 65 (2014) 161–175. doi:10.1016/j.actamat.2013.10.057.

[10] H. Wang, S. Patala, B. J. Reich, Constrained bayesian nonparametric regression for grain boundary energy predictions, 2020. URL: https://arxiv.org/abs/2002.00938. doi:10.48550/ARXIV.2002.00938.

[11] S. G. Baird, E. R. Homer, D. T. Fullwood, O. K. Johnson, Five degree-of-freedom property interpolation of arbitrary grain boundaries via voronoi fundamental zone framework, Computational Materials Science 200 (2021) 110756. doi:10.1016/j.commatsci.2021.110756.

[12] O. K. Johnson, J. M. Lund, T. R. Critchfield, Spectral graph theory for characterization and homogenization of grain boundary networks, Acta Materialia 146 (2018) 42–54. doi:10.1016/j.actamat.2017.11.054.

[13] A. Lecchini-Visintini, J. Lygeros, J. Maclejowski, Simulated annealing: Rigorous finite-time guarantees for optimization on continuous domains, Advances in Neural Information Processing Systems 20 - Proceedings of the 2007 Conference (2009).

[14] W. L. Hamilton, R. Ying, J. Leskovec, Representation learning on graphs: Methods and applications, 2017. URL: https://arxiv.org/abs/1709.05584. doi:10.48550/ARXIV.1709.05584.

[15] J. L. Priedeman, C. W. Rosenbrock, O. K. Johnson, E. R. Homer, Quantifying and connecting atomic and crystallographic grain boundary structure using local environment representation and dimensionality reduction techniques, Acta Materialia 161 (2018) 431–443. doi:10.1016/j.actamat.2018.09.011.

[16] M. R. Lemes, L. R. Marim, A. Dal Pino, Application of artificial intelligence to search ground-state geometry of clusters, Physical Review A - Atomic, Molecular, and Optical Physics 66 (2002) 5. doi:10.1103/PhysRevA.66.023203.

[17] E. P. P. Pe-Than, D. H. L. Goh, C. S. Lee, A typology of human computation games: An analysis and a review of current games, Behaviour and Information Technology 34 (2015) 809–824. doi:10.1080/0144929X.2013.862304.

[18] O. Tremblay-Savard, A. Butyaev, J. Waldispuhl, Collaborative solving in a human computing game using a market, skills and challenges, CHI PLAY 2016 - Proceedings of

the 2016 Annual Symposium on Computer-Human Interaction in Play (2016) 130–141. doi:10.1145/2967934.2968104.

[19] M. Ponti, I. Stankovic, W. Barendregt, B. Kestemont, L. Bain, Chefs know more than just recipes: Professional vision in a citizen science game, Human Computation 5 (2018) 1–12. URL: https://hcjournal.org/index.php/jhc/article/view/88. doi:10.15346/hc.v5i1.1.

[20] F. Khatib, F. Dimaio, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska, I. Pichova, J. Thompson, Z. Popović, M. Jaskolski, D. Baker, Crystal structure of a monomeric retroviral protease solved by protein folding game players, Nature Structural and Molecular Biology 18 (2010) 1175–1177. doi:10.1038/nsmb.2119.

[21] D. Sels, Stochastic gradient ascent outperforms gamers in the quantum moves game, Physical Review A 97 (2018) 1–6. doi:10.1103/PhysRevA.97.040302.

[22] N. R. Prestopnik, K. Crowston, Gaming for (citizen) science: Exploring motivation and data quality in the context of crowdsourced science through the design and evaluation of a social-computational system, Proceedings - 7th IEEE International Conference on e-Science Workshops, eScienceW 2011 (2011) 28–33. doi:10.1109/eScienceW.2011.14.

[23] N. Prestopnik, K. Crowston, J. Wang, Exploring Data Quality in Games With a Purpose, in: iConference, 2014. doi:10.9776/14066.

[24] B. Morschheuser, J. Hamari, J. Koivisto, A. Maedche, Gamified crowdsourcing: Conceptualization, literature review, and future agenda, International Journal of Human Computer Studies 106 (2017) 26–43. doi:10.1016/j.ijhcs.2017.04.005.

[25] R. M. Ryan, C. S. Rigby, A. Przybylski, The motivational pull of video games: A self-determination theory approach, Motivation and Emotion 30 (2006) 347–363. doi:10.1007/s11031-006-9051-8.

[26] L. von Ahn, L. Dabbish, Designing games with a purpose, Communications of the ACM 51 (2008) 57. doi:10.1145/1378704.1378719.

[27] M. C. Yuen, L. J. Chen, I. King, A survey of human computation systems, Proceedings - 12th IEEE International Conference on Computational Science and Engineering, CSE 2009 4 (2009) 723–728. doi:10.1109/CSE.2009.395.

[28] A. Kawrykow, G. Roumanis, A. Kam, D. Kwak, C. Leung, C. Wu, E. Zarour, L. Sarmenta, M. Blanchette, J. Waldispühl, Phylo: A citizen science approach for improving multiple sequence alignment, PLoS ONE 7 (2012). doi:10.1371/journal.pone.0031362.

[29] K. Tuite, Gwaps: Games with a problem, Foundations of Digital Games 2014 (2014).

[30] N. R. Prestopnik, J. Tang, Points, stories, worlds, and diegesis: Comparing player experiences in two citizen science games, Computers in Human Behavior 52 (2015) 492–506. doi:10.1016/j.chb.2015.05.051.

[31] N. Lane, N. R. Prestopnik, Diegetic connectivity: Blending work and play with storytelling in serious games, Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '17 (2017) 229–240. doi:10.1145/3116595.3116630.

[32] N. Prestopnik, K. Crowston, J. Wang, Computers in human behavior gamers , citizen scientists , and data : Exploring participant contributions in two games with a purpose, Computers in Human Behavior 68 (2017) 254–268. doi:10.1016/j.chb.2016.11.035.

[33] R. Tinati, M. Luczak-Roesch, E. Simperl, W. Hall, An investigation of player motivations in eyewire, a gamified citizen science project, Computers in Human Behavior 73 (2017) 527–540. doi:10.1016/j.chb.2016.12.074.

[34] A. J. Quinn, B. B. Bederson, Human computation: A survey and taxonomy of a growing field, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2011) 1403–1412. doi:10.1145/1978942.1979148.

[35] K. Siu, M. O. Riedl, Reward systems in human computation games, in: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 266–275. URL: https://doi.org/10.1145/2967934.2968083. doi:10.1145/2967934.2968083.

[36] G. Li, Human-in-the-loop data integration, Proceedings of the VLDB Endowment 10 (2017) 2006–2017. doi:10.14778/3137765.3137833.

[37] M. Jia, Y. Wei, Y. Ma, Y. Shen, Y. Zhang, A case study of expressively constrainable level design, Zhongguo Zhong Yao Za Zhi 22 (1997) 454–458,510.

[38] M. T. Cheng, L. Rosenheck, C. Y. Lin, E. Klopfer, Analyzing gameplay data to inform feedback loops in the radix endeavor, Computers and Education 111 (2017) 60–73. doi:10.1016/j.compedu.2017.03.015.

[39] L. Von Ahn, M. Blum, N. J. Hopper, J. Langford, Captcha: Using hard ai problems for security, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2656 (2003) 294–311. doi:10.1007/3-540-39200-9_18.

[40] D. Gundry, S. Deterding, Validity threats in quantitative data collection with games: A narrative survey, Simulation and Gaming (2018). doi:10.1177/1046878118805515.

[41] E. Andersen, E. O'Rourke, Y.-E. Liu, R. Snider, J. Lowdermilk, D. Truong, S. Cooper, Z. Popovic, The impact of tutorials on games of varying complexity, Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12 (2012) 59. doi:10.1145/2207676.2207687.

[42] K. Siu, A. Zook, M. O. Riedl, Collaboration versus competition: Design and evaluation of mechanics for games with a purpose, Proceedings of the 9th International Conference on the Foundations of Digital Games (FDG '14) (2014).

[43] K. Siu, M. Guzdial, M. O. Riedl, Evaluating singleplayer and multiplayer in human computation games, in: Proceedings of the 12th International Conference on the Foundations of Digital Games, FDG '17, Association for Computing Machinery, New York, NY, USA, 2017. URL: https://doi.org/10.1145/3102071.3102077. doi:10.1145/3102071.3102077.

[44] S. Ahmad, A. Bryant, E. Kleinman, Z. Teng, T. H. D. Nguyen, M. S. El-Nasr, Modeling individual and team behavior through spatio-temporal analysis, CHI PLAY 2019 - Proceedings of the Annual Symposium on Computer-Human Interaction in Play (2019) 601–612. doi:10.1145/3311350.3347188.

[45] J. Gaston, S. Cooper, To three or not to three: Improving human computation game onboarding with a three-star system, Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17 (2017) 5034–5039. doi:10.1145/3025453.3025997.

[46] E. Simperl, N. Reeves, C. Phethean, T. Lynes, R. Tinati, 6 is virtual citizen science a game?, ACM Trans. Soc. Comput 1 (2018). doi:10.1145/3209960.

[47] S. Horowitz, B. Koepnick, R. Martin, A. Tymieniecki, A. A. Winburn, S. Cooper, J. Flatten, D. S. Rogawski, N. M. Koropatkin, T. T. Hailu, N. Jain, P. Koldewey, L. S. Ahlstrom, M. R. Chapman, A. P. Sikkema, M. A. Skiba, F. P. Maloney, F. R. Beinlich, Z. Popovic, D. Baker, F. Khatib, J. C. Bardwell, Determining crystal structures

through crowdsourcing and coursework, Nature Communications 7 (2016). doi:10.1038/ncomms12549.

[48] M. Krause, J. Smeddinck, Human computation games: A survey, European Signal Processing Conference (2011) 754–758.

[49] Little big planet, 2008.

[50] Minecraft, 2013.

[51] R. W. Balluffi, S. M. Allen, W. C. Carter, Kinetics of materials, Wiley-Interscience, 2005. doi:10.1002/0471749311.

[52] H. Grimmer, Disorientations and coincidence rotations for cubic lattices, Acta Crystallographica Section A 30 (1974) 685–688. URL: https://doi.org/10.1107/S0567739474001719. doi:10.1107/S0567739474001719.

[53] H. Grimmer, A unique description of the relative orientation of neighbouring grains, Acta Crystallographica Section A 36 (1980) 382–389. URL: https://doi.org/10.1107/S0567739480000861. doi:10.1107/S0567739480000861.

[54] R. Quey, P. R. Dawson, F. Barbe, Large-scale 3d random polycrystals for the finite element method: Generation, meshing and remeshing, Computer Methods in Applied Mechanics and Engineering 200 (2011) 1729–1745. doi:10.1016/j.cma.2011.01.002.

[55] B. Mohar, The laplacian spectrum of graphs, Graph Theory, Combinatorics, and Applications 2 (1991) 871–898. doi:10.1017/CBO9781107415324.004.

[56] D. Knossow, A. Sharma, D. Mateus, R. Horaud, Inexact matching of large and sparse graphs using laplacian eigenvectors, in: A. Torsello, F. Escolano, L. Brun (Eds.), Graph-Based Representations in Pattern Recognition, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 144–153.

[57] O. Šedivý, T. Brereton, D. Westhoff, L. Polívka, V. Beneš, V. Schmidt, A. Jäger, 3d reconstruction of grains in polycrystalline materials using a tessellation model with curved grain boundaries, Philosophical Magazine 96 (2016) 1926–1949. doi:10.1080/14786435.2016.1183829.

[58] J. Lazar, J. Feng, H. Hochheiser, Research methods in human-computer interaction, 2 ed., Morgan Kaufmann, Oxford, England, 2017.

[59] D. Manova, S. Mändl, H. Neumann, B. Rauschenbach, Influence of grain size on nitrogen diffusivity in austenitic stainless steel, Surface and Coatings Technology 201 (2007) 6686–6689. doi:10.1016/j.surfcoat.2006.09.104.

[60] I. Yildirim, Time pressure as video game design element and basic need satisfaction, CHI Extended Abstracts on Human Factors in Computing Systems (2016) 2005–2011. doi:10.13140/RG.2.1.4107.0564.