The effect of performance feedback and optimization on the conceptual design process

Nathan C. BROWN[1*], Caitlin T. MUELLER[1]

[1]Massachusetts Institute of Technology
77 Massachusetts Ave, Room 5-418
Cambridge, MA 02143
*ncbrown@mit.edu

Abstract

Many recent contributions in computational structural design have argued that design quality can be improved when performance feedback and guidance are part of the conceptual design process. However, the effect of multi-objective feedback and guidance tools has not been studied extensively. This paper presents the results of an educational study that tests the direct relationship between conceptual design tools and the simulated performance of resulting designs. In the study, students were tasked with designing a restaurant canopy roof using a series of increasingly performance-driven computational design tools. Although there was no consensus on preferred workflows or aesthetic preferences, the average designs chosen using real time feedback or directed optimization performed significantly better in terms of deflection and emissions than those chosen through free exploration. Overall, this research establishes a link between design tools and performance outcomes, while strengthening the argument for further integration of performance feedback into early stage design processes.

Keywords: conceptual design, design tradeoffs, interactive design, performance feedback, behavioral study

1. Introduction

In recent years, researchers in the area of computational design for structures have emphasized the development of tools that enable both performance feedback and guidance in early stage design. These contributions respond to limitations of the traditional design process in which design and analysis software are entirely separate, making substantial design iterations that respond to performance difficult and time-consuming to pursue. As a result, engineers and other specialists are often given a rigid design and limited to small adjustments or simply "making it work". Researchers assert that when performance analysis is present in the conceptual phase, which is when a design is most flexible, it has much greater potential to improve the performance of an eventual building or structure. Due to its context of structural design, much of this research is concerned with large, global decisions such as typology, material, or specific geometry that are difficult to change later in the design process, but typically have a large influence on building performance.

This paper tests the assertion of improved performance by investigating the effect of performance feedback and optimization techniques on architectural design processes through a behavioral case study. In the study, a group of design students are given a structural design problem and provided with a series of increasingly performance-driven computational design tools to complete the task. The designs chosen by the students are then analyzed for quality and diversity while comparing outcomes across different design environments. Overall, this research seeks a better understanding of the relationship between performance data and design outcomes in early stage design.

2. Literature Review

Design studies, which focuses on developing an understanding of design processes across areas such as engineering, product, architectural, and urban design, is an established academic field with robust supporting literature. An initial contribution towards establishing the field is given by Cross *et al.* [2], which argues that design should be regarded as a technology rather than a science, since design and technology include the application of knowledge other than the purely 'scientific' kind. A thorough review of early design theory and methodology in mechanical engineering is provided in Finger & Dixon [4]. Within the examination of design processes, behavioral studies that test the relationship between different engineering design environments and idea generation are common. Shah *et al.* [12] establish experimental guidelines for evaluating conceptual design strategies, and these guidelines are widely followed. Mckoy *et al.* [7] analyze the influence of design representation on the effectiveness of idea representation. Schlecht [11] tests the impact of prototyping environments on ideation, while Faas *et al.* [3] address the question of whether or not designers who are more engaged, as measured by presence and immersive tendency questionnaires, produce better designs.

Architectural design provides its own developing history of design processes, which have recently shifted from drawings and physical modeling to computational, generative, and performance-based parametric and morphological models (Oxman [9]). The requirements of an architectural design and subsequent decision-making processes are often subjective and particular to the field, which can complicate efforts to evaluating the quality and novelty of designs using traditional means. While specific architectural design processes have been studied from a behavioral standpoint (Suwa & Tversky [14], Goldschmidt [5]), there is a clear need for research contributions that explore how designers interact with newer digital design methodologies. The recent emphasis on high-performance architecture and the integration of analysis tools into conceptual design workflows demands additional testing on how these feedback mechanisms influence the design process. Many recent contributions to the computational design of structures are based on the assumption that overall design quality can be improved when performance feedback and guidance are part of the conceptual design process (Mueller & Ochsendorf [8]). Some efforts have been made to test this theory, including by Arnaud [1]. However, many architectural and even structural design problems are multi-objective, requiring the synthesis of feedback in multiple dimensions. This paper proposes a behavioral study in which participants must pursue multiple objectives simultaneously, with the intent of uncovering more generalizable knowledge concerning the relationship between performance input and quality in conceptual design.

3. Methodology

3.1 Participants

The study involved 26 undergraduate and graduate students at a U.S. University. The vast majority of participants were in their second year of a graduate architectural degree program, while 3 were pursuing degrees in civil and environmental engineering. All of these students were taking an architecture class in building structural systems and were given the design task as an educational exercise. Students in the class were provided the option of participating or requesting that their chosen designs not be included in the aggregate study results. Materials including software files and instructions were distributed online, and participants were able to complete the exercise on their own computers and upload answers at their convenience. Most participants had at least a year worth of formal design training, including two structures courses, and were concurrently taking a comprehensive architectural design studio.

3.2. Procedure

Each experiment consisted of three design phases plus a short survey about the overall experience. The phases consisted of the same design problem, but with different conditions that progressively gave

participants more access to performance-driven computational tools. In the first phase, students were given a design problem and a corresponding parametric model that defined the potential geometry of the architectural solution. Participants were free to adjust any of the design variables while the corresponding geometry updated in real time. The second phase consisted of the same parametric model and variable sliders, but with performance feedback also updating on the screen along with the model geometry. This feedback included both actual and normalized performance values for a variety of design objectives, as well a simple bar graph to visualize the performance scores. In the third phase, participants were instructed to use a number of optimization tools to guide their design exploration. The students were asked to record their "favorite" design along with 5 desirable alternatives for each phase. After completing the exercise, participants were given survey questions comparing the three different design environments in terms of ease of use, quality of outcomes, likeliness of using in their own workflows, and related topics. Participants were provided with a prepared form to record all of their answers.

3.3. Task

The task given in this study was the design of a canopy structure for the outdoor seating area of the restaurant (Figure 1). Due to the desire for a free edge and the ability to anchor into the wall above, the hypothetical client asked for a cable-stayed structure. The main topology of the structure is formed by beams that cantilever out from the wall and are supported by a series of cables, which also anchor into the wall. Within this main geometry, participants were allowed to adjust the anchor point spread, height of cable and beam connections, height and horizontal distance to the canopy tip, number of cables, curvature of the canopy, and the structural material. A full description of the adjustable design variables is given in Table 1.
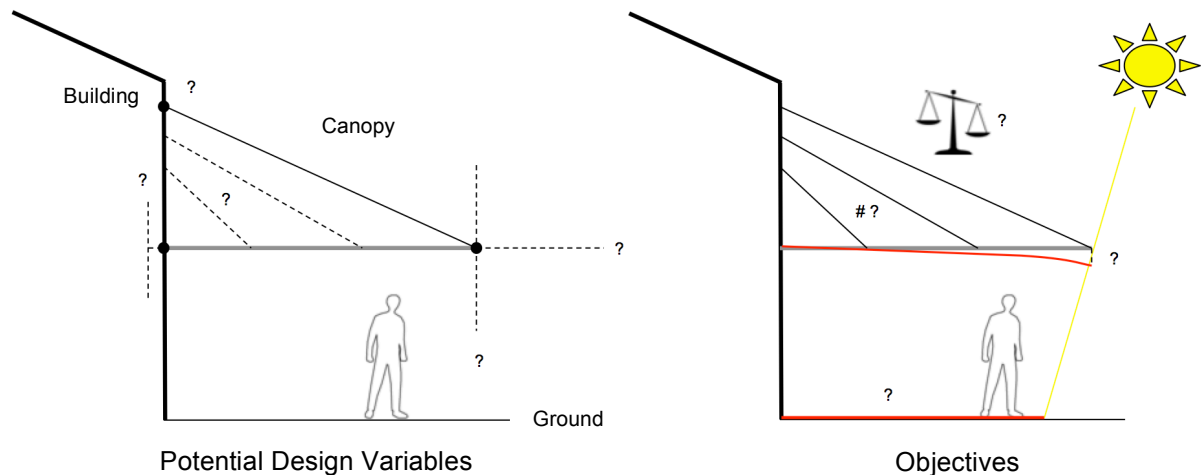

Figure 1: The design problem, a cable-stayed canopy roof, as presented to participants

As part of the exercise, a number of computational models were built to quantitatively assess the performance of each design. These models calculated the shaded area provided, carbon emissions due to the materials, number of connections, and maximum deflection of the design. The shaded area provided is an architectural metric that would be a priority of the client, and it was calculated geometrically using an assumed static sun angle. The number of connections roughly approximates how difficult the design is to construct. Carbon emissions and maximum deflection are structural metrics and were calculated using finite element analysis. Carbon emissions is linked to material selection and quantity, while maximum deflection indicates response to loads. Table 2 gives additional information about the objectives and their evaluation methods.

Table 1: Design space variables and bounds for the design task given to participants

Symbol	Variable	Range
x_1	Anchor Point Spread	$0 < x < 1$
x_2	Height of (Top) Cable Anchor Point	$2.44 \text{ m} < x < 9.14 \text{ m}$
x_3	Height of Canopy Anchor Point	$2.13 \text{ m} < x < 7.62 \text{ m}$
x_4	Length of Canopy	$1.52 \text{ m} < x < 12.19 \text{ m}$
x_5	Height of Canopy Tip	$2.13 \text{ m} < x < 4.57 \text{ m}$
x_6	Number of Cables	$1 < x < 10$
x_7	Curvature	$0 < x < 1$
x_8	Material	Steel, Aluminum, Wood, Carbon Fiber

The stated goal of the design exercise, which is implicit in most design situations, was to prioritize, navigate, and explore interrelated performance and aesthetic objectives to arrive at a satisfying design solution. Participants were given the freedom to choose which objectives were the most important and encouraged to use their own design sensibilities when judging the expressiveness of the solution. Each objective was normalized so that the best performing design received a score of '1', and every other score is a multiple showing how much worse the design performs than the optimal. Both raw and normalized feedback were given to the participant, while the normalized values were used in the optimization parts of the exercise.

Table 2: Quantitative objective functions for the design experiment

Objective	Metric	Evaluation Method
Shade	Area shaded by canopy (max, ft^2)	Geometric measurement; assuming 50° sun angle
Connections	Number of connections (min)	Sum of intersections between cables and canopy/wall
Carbon Emissions	Emissions due to material volume (max, kg CO_2)	FEM + Sizer
Deflection	Maximum deflection in model (min, in)	FEM

3.4. Materials

Each participant was provided with a parametric model of the design space created by the author. The model was developed in Rhinoceros and Grasshopper. The model geometry and the evaluation metrics for shade and connections were built using native components. The structural model was created using Karamba [10], a plug-in for Grasshopper that interacts with Rhinoceros geometry. A vertical distributed load of 0.60 kN/m was applied to each beam, which corresponds to the tributary area of multiple beams arrayed longitudinally along the wall of the restaurant. To estimate the overall weight of structural material, Karamba's sizing feature was used. This feature checks the allowable axial, bending, and buckling loads for each member and then searches through a structural section library to determine the smallest member that can adequately handle each load before outputting the total weight of the structural system. These material quantities were then multiplied by carbon coefficients calculated in the Inventory of Carbon & Energy database (Hammond & Jones [6]) to arrive at emissions.

Although a more detailed analysis involving uplift and other forces could be completed on the structure, this study focuses on conceptual design, and thus the evaluations are greatly simplified so that designs can be explored instantaneously and compared based on relative performance. The purpose was to give participants rapid feedback to build intuition about the design space and ultimately guide exploration while maintaining creative flow. For this reason, square footage and

deflection were given in Imperial units, since these units might be understood more quickly by students with experience in U.S. universities and design offices. Units for emissions were given in kg of CO_2, which is consistent with the ICE Database. In addition to the model shown in Figure 2, a 3D version was created to allow for creative exploration of curvature as part of the educational exercise. However, the results of this paper only include the 2D model.

In phase 1 of the exercise, only the geometry of the design was shown to participants. In phase 2, this same geometry was shown alongside the performance feedback information. In phase 3, students were able to use either Galapagos, an evolutionary solver native to Grasshopper, or one of the derivative-free optimization algorithms contained within Goat, which is an alternative optimization plug-in. These optimization components were connected to a composite function combining each of the objective scores along with an adjustable importance weight for each objective. Although participants were encouraged to explore different weight combinations as part of the exercise, the instructions were to simply use the tools to produce satisfying designs, matching the task of the earlier two phases.
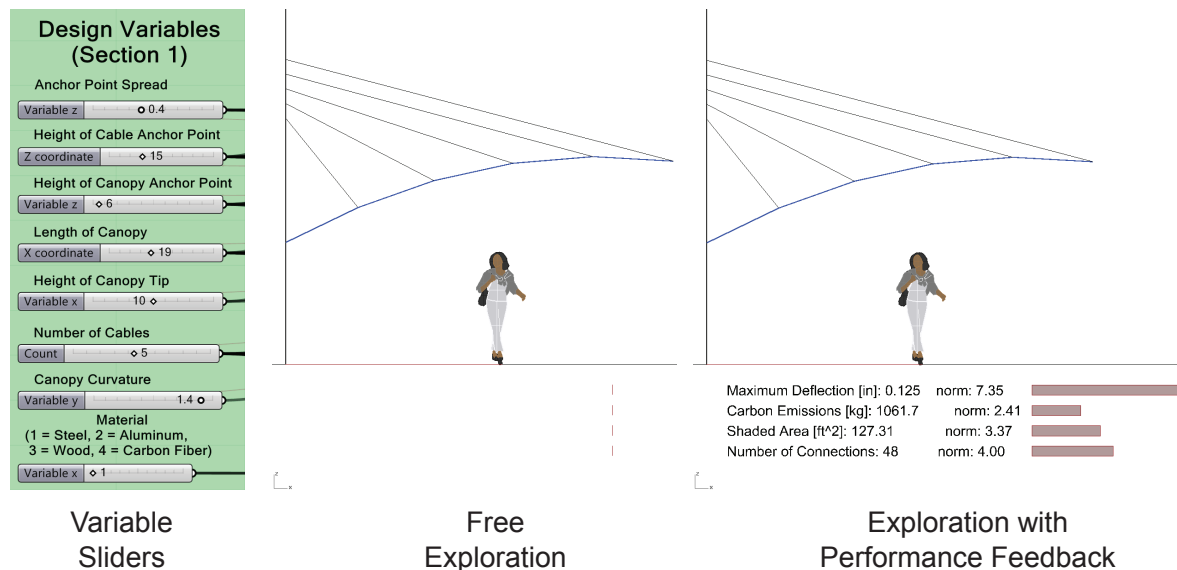

Figure 2: The main parametric model and design environment, showing phases 1 and 2

3.5. Design outcome measures: quality and diversity

The results of this study include both the measured quality of the designs produced by the participants and an evaluation of their novelty and diversity. The quality of the designs refers to their relative performance in each of the four dimensions. Performance data collected throughout the study was separated by both phase and objective. Outliers that did not comply with the original bounds of the problem were removed from each dataset. A number of statistical tests were then completed on the datasets to compare overall performance averages and medians between the three different design environments. To determine any significant effect the performance feedback and optimization tools had on design quality, a single factor ANOVA test was completed on the three datasets for each performance measurement. This analysis tests the null hypothesis that the averages of each set are statistically equal. If the ANOVA test determined that at least one of the datasets was significantly different, a series of two-tail t-tests were completed on the different sets to find out which of the performance categories were statistically different from the free exploration dataset. Each of these tests used an alpha of 0.05 for a 95% confidence level.

In addition to the quality of examples produced in a conceptual design exercise, researchers are also concerned with creativity, and there are established criteria for assessing this aspect of a particular design environment. These criteria include novelty and variety, which can both be obtained by defining what is not novel, or by identifying key attributes and functionalities, developing a hierarchical rating system, and scoring each design based on this system (Shah *et al.* [13]). It is also possible to mathematically compute the diversity of a dataset, which measures how different the designs in a set are from one another. This diversity measurement can be completed by computing certain geometric relationships between the different design vectors, such as the radius of the smallest enclosing ball, area or perimeter of a convex hull, or taking an average or total distance to the centroid of all design points in n-dimensional space, where n is the number of design variables.

Novelty, variety, and diversity are all important concepts in this exercise due to the creative requirements of architectural and structural design processes. If a design tool produces high-performing designs but constrains results to visually uninteresting solutions or only a small portion of the design space, it is likely not useful to expressive designers desiring the ability to exert preference. However, the most common metrics for novelty and variety both require researchers to dictate a scoring system for the design problem, which is unsuitable for the open-ended nature of this study. The author ran a number of different diversity calculations on each dataset, but the metrics did not agree in their ranking and comparison between the different design environments. Consequently, visualizations of the design vectors and example geometries generated for each phase will be presented directly in the results section of this paper, along with commentary on their meaning.

3.6 Participant survey

Upon finishing the exercise, participants were asked to complete a short survey related to their experience. The survey asked participants to rank the design environments (free exploration, performance feedback, optimization tools) on a number of factors: which environment is the easiest to work with, leads to the best design outcome, they would most like to mention to a client or architectural reviewer, and they are most likely to apply in their own design workflows. Analysis of these responses supplements the measurement of design quality and discussion of design diversity.

4. Results and discussion

4.1 Design quality

The aggregate performance scores for each experiment are given in the box plots in Figure 3. The plots show the median, interquartile range, upper and lower whiskers, and outliers for each dataset. A note is also included if a dataset for performance feedback or optimization tools differed significantly from the dataset for free exploration. In the shade performance metric, only the optimization tools dataset showed any significant difference, with designs in this set performing substantially worse than in free exploration. There was no statistical difference in the number of connections measured between the design environments. Both performance feedback and optimization tools significantly reduced the estimated embodied carbon in selected designs compared to free exploration. While the median deflection values for both feedback and optimization were lower than in free exploration, only the optimization dataset registered as significantly lower.

The lack of improvement in shaded length and number of connections can be partially attributed to the fact that these objectives are easy to understand visually regardless of design environment. In addition, these performance metrics seem not to have been prioritized as highly as the structural metrics across all three phases. This effect could be a result of the study's educational setting, or of the differences in magnitude between normalized scores for lower performers—carbon and deflection could score hundreds or thousands of times worse than the optimal, while the design space and normalization technique only allowed for much smaller scores in shade and connections.
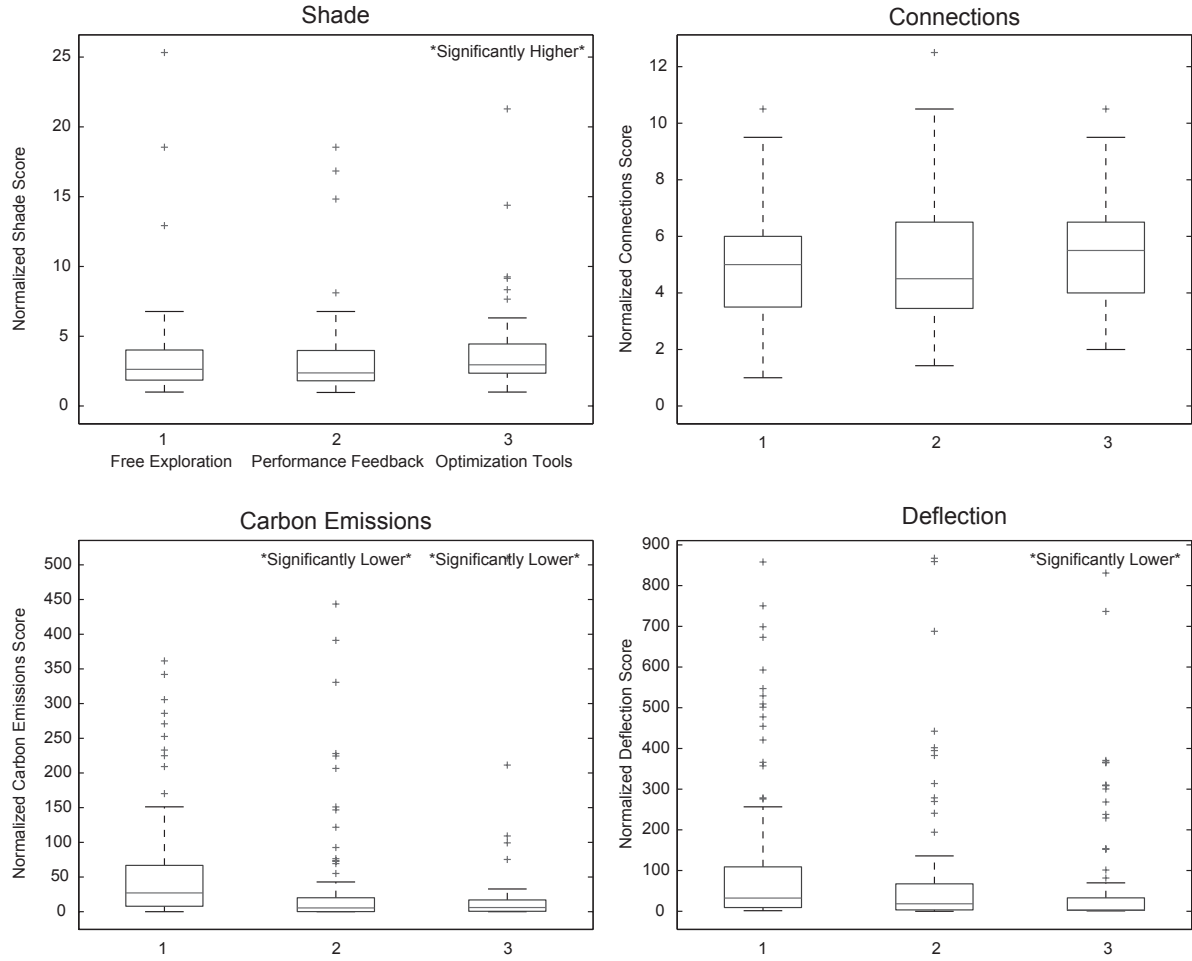
Figure 3: Box plots showing performance scores of designs for different environments

In contrast, the performance scores of embodied carbon and deflection improved significantly when designers were given additional data. The designs generated while using performance feedback and optimization tools resulted in 43% and 68% lower carbon emissions on average, respectively. Structural performance is less easy to predict intuitively unless designers have advanced education or experience in the field. The results of this study suggest that access to computational tools helps mitigate lack of structural intuition, while guiding participants towards better performing designs. Optimization tools magnified this effect, since participants seemed more willing to balance the four objectives during exploration with feedback than during optimization, which pushed priorities towards the structural criteria at the significant expense of shaded length. The optimization tools also reduced the number of very poorly performing outliers, mostly because optimization is unlikely to pick a design solution with large scores in any category unless that category is completely ignored in the composite function.

4.2 Design diversity

All of the design vectors produced during the exercise are displayed in the parallel coordinate plots in Figure 4. A representative sample of designs produced in each environment is also shown geometrically in Figure 5. Each cleaned dataset contained over 110 designs, making it impossible to show every single result. In all three cases, the vector plots leave few gaps, illustrating that participants covered a large portion of the design space and did not produce designs that can be

categorized simply. Nevertheless, a number of trends are visible when viewing both the design vector plots and representative examples. One noticeable difference between environments is that fewer carbon fiber and aluminum solutions are present when either feedback or optimization is utilized. Both of these materials are considered high-performance by many in the architectural community due to their strength and weight, leading to numerous generated solutions during free exploration. However, these materials have energy-intensive manufacturing practices and relatively high carbon coefficients, which becomes evident once performance feedback is enabled. The vector plots also show that the optimization tools favored designs with less extreme values for number of cables and curvature of the canopy, which corresponds to subtle curves and more than one or two cables.
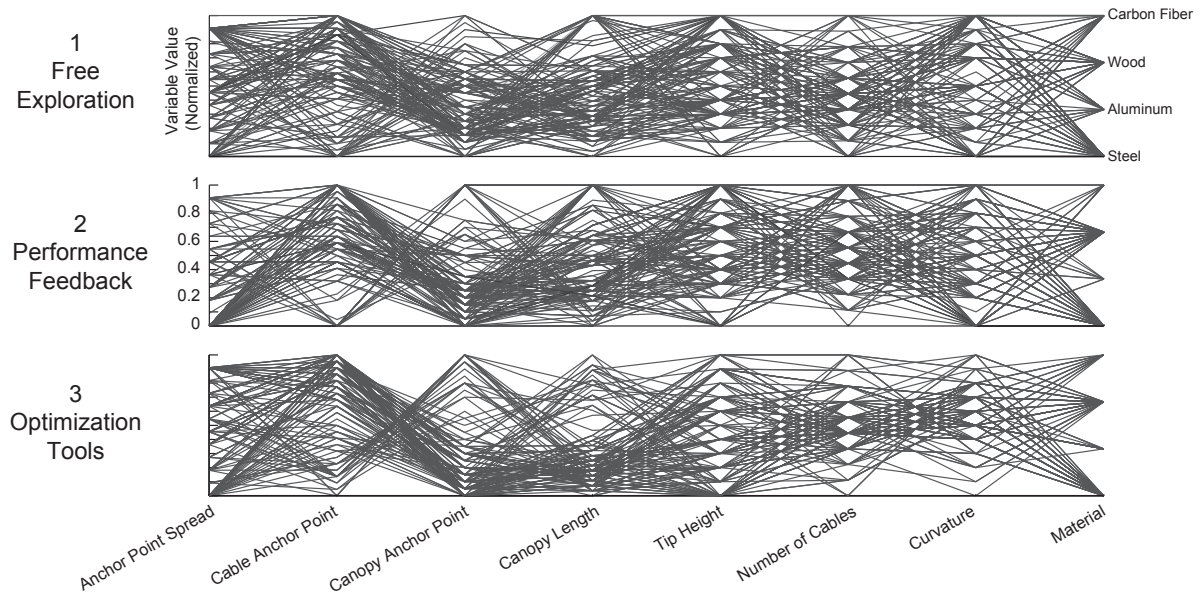
Figure 4: Parallel coordinate plots of the design vectors for each collected dataset
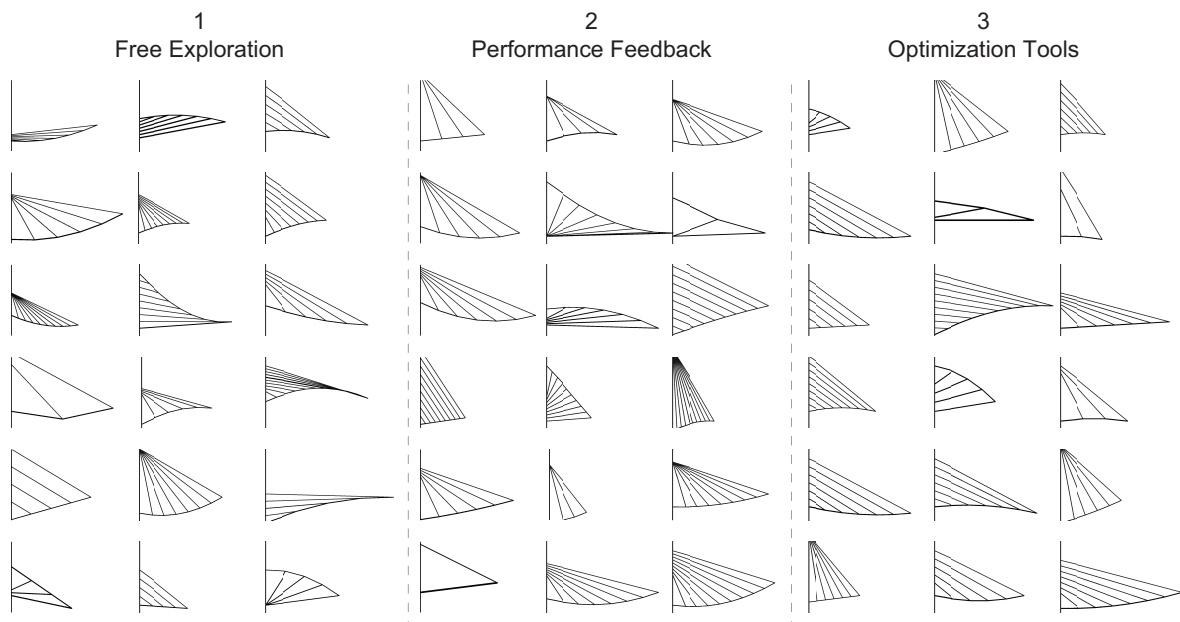
Figure 5: Representative designs generated during each phase of the study

In addition, the optimization tools pushed designs to have lower tip heights, which improves shaded area without a large negative impact on other metrics. This effect is somewhat noticeable for the performance feedback phase, but it is especially obvious when comparing the prevalence of designs with the lowest tip height in free exploration versus optimization tools. Similarly, structural performance feedback and optimization tools encouraged designers to produce geometries with low canopy anchor points and high cable points, as indicated by the high density of lines moving diagonally between the two extreme values. Interaction with performance data also drove canopy anchor heights much lower in general. A number of creative participants realized that setting a canopy anchor point above the cables turned them into compressive struts and yielded a different typology, but these designs required considerably more material to resist buckling and also generally paid a penalty on shaded area due to high tip heights, thus making them undesirable in performance-based design. Despite these noticeable differences, there was still a high degree of geometric diversity present in each design environment. Although performance feedback and optimization both discouraged some areas of the parametric design space, the large variations demonstrated in Figure 5 suggest that designers still had sufficient flexibility to exert preference.

4.3. Survey Responses

After completing the exercise, students were asked a series of optional questions related to the three different design environments. The results of their answers are given in Table 3. None of the questions led to an overwhelming winner in any of the categories. Surprisingly, the single most common response was that optimization tools were the easiest to manage, perhaps because participants felt that the algorithms were doing part of the work of eliminating bad outcomes. Although participants seemed to feel on average that performance data does improve design quality, respondents were perfectly split on which technique would give them the most confidence when explaining their design to a client or critic. Overall, it is clear that different designers have different preferences, and flexibility is often key when setting up a design environment.

Table 3: A comparison of participant preferences for the three different design environments

Question	Free Exploration	Performance Feedback	Optimization Tools
Easiest to work with:	30%	15%	55%
Leads to best design outcome:	24%	38%	38%
Would most like to mention to a client or reviewer:	33%	33%	33%
Most likely to apply in their own workflow:	24%	52%	24%

5. Conclusion: summary of contributions

In conclusion, this paper presents new data concerning the effects of performance feedback and guidance on the conceptual design of structures through an experimental case study. The study, composed of students with formal design training, gave participants the task of designing a cable-stayed canopy roof, along with a parametric model for design exploration and progressive access to performance-based computational tools. The design exercise contained three phases, which each asked the students to record their favorite designs. Participants were then asked to assess their experience in each of the different environments.

Although there was no consensus between the students on aesthetic preferences, prioritization of design variables and objectives, or favorite design workflow, designs chosen using either real-time performance feedback or a directed optimization process performed significantly better in terms of deflection and emissions than designs chosen through free exploration. The average effect of feedback and guidance on the shaded area and connections metrics, which can be more easily

determined visually, was not significant except in the case of shaded area, which was made worse when students used optimization. In terms of design diversity, performance feedback and optimization tools virtually eliminated some poor-performing areas of the design space, while encouraging a number of specific design characteristics, such as low canopy tip heights and a large spread between cable and canopy anchors. However, there was still noticeable variation and diversity within the designs produced using performance-based methods, which indicates that flexibility and creativity were still possible even as considerations of performance clearly guided participants. Overall, this research determined that the use of performance data in conceptual design can have a noticeably positive effect, a finding that encourages the further development of performance-based computational tools. Future study in this area could include additional case studies, a higher number of participants, and more controlled settings. However, this research is an initial quantitative step in testing and supporting the arguments commonly made for greater integration of rapid, multi-dimensional performance feedback into architectural and structural conceptual design workflows.

References

[1] Arnaud V.B., Quantifying Architects' and Engineers' Use of Structural Design Software, M.Eng. Thesis, Massachusetts Institute of Technology, 2013.

[2] Cross N., Naughton J. and Walker D., Design method and scientific method, *Design Studies*, 1981; **2(4)**; 195–201.

[3] Faas D., Bao Q., Frey D.D. and Yang M.C., The influence of immersion and presence in early stage engineering designing and building, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 2014; **2**; 139–151.

[4] Finger S. and Dixon J.R., A review of research in mechanical engineering design. part I: descriptive, prescriptive, and computer-based models of design processes, *Research in Engineering Design*, 1989; **1**; pp. 51–67, 1989.

[5] Goldschmidt G., On visual design thinking: the vis kids of architecture, *Design Studies*, 1994; **15(2)**; 158–174.

[6] Hammond G. and Jones C., Inventory of Carbon & Energy (ICE), Version 1.6a., University of Bath, 2010.

[7] Mckoy F.L., Vargas-Hernández N., Summers J.D., and Shah J.J., Influence of design representation on effectiveness of idea generation, in *Proceedings of DETC'01: ASME 2001 Design Engineering Technical Conferences*, 2001, 1–10.

[8] Mueller C.T. and Ochsendorf J.A., Combining structural performance and designer preferences in evolutionary design space exploration, *Automation in Construction*, 2015; **52**; 70–82.

[9] Oxman R., Digital architecture as a challenge for design pedagogy: theory, knowledge, models and medium, *Design Studies*, 2008; **29**; 99–120.

[10] Preisinger C. and Bollinger-Grohmann-Schneider ZT GmbH, Karamba. Bollinger-Grohmann-Schneider ZT GmbH, Vienna.

[11] Schlecht L.A., Impact of prototyping resource environments on idea generation in product design, S.M. Thesis, Massachusetts Institute of Technology, 2013.

[12] Shah J.J., Kulkarni S. and Vargas-Hernandez N., Guidelines for experimental evaluation of idea generation methods in conceptual design, *Journal of Mechanical Design*, 2000; **122(4)**; 337–384.

[13] Shah, J.J., Vargas-Hernande N, and Smith S.M., Metrics for measuring ideation effectiveness, *Design Studies,* 2003; **24**; 111–134.

[14] Suwa M. and Tversky B., What do architects and students perceive in their design sketches? A protocol analysis, *Design Studies*, 1997; **18**; 385–403.