Critical Review: Improving the Efficacy of Games for Change Using Personalization Models

Introduction

Persuasion is the act of influencing or reinforcing certain attitudes and behaviours [Khaled et al., 2008] and use of technology for this purpose to benefit users or the wider community is a long-term area of research [Fogg, 2002]. This review summarises and critically analyses Orji et al. [2017] whose paper explores the use of personalisation models as a persuasive device for improving the efficacy of games designed to change user behaviour, to determine future work in this domain of Persuasive Technology.

Summary of Contributions

Orji et al. [2017] discuss the rising prominence of *games for change*[1] designed for purposes other than entertainment, to educate players about topics in a way that influences their behaviour [Busch et al., 2015]. Many such games are designed with a problematic "*one size fits all*" philosophy - the game design (i.e. its adopted persuasive *strategies*) are not tailored to the *type* of player. Orji et al. [2017] therefore seek to answer two main research questions to understand if the efficacy of certain strategies in existing games [Peng, 2009, Kaipainen et al., 2012] can be maximised by catering to the *type* of player. Firstly, whether tailoring games for change to a specific player type increases their persuasiveness. Secondly, if beneficial effects of tailoring are indeed observed, whether these are mediated by an improved playing experience. By answering these, results may inform decisions of games designers as to which persuasive strategies they adopt to maximise efficacy in certain player types.

Treating user groups in a monolithic way is generally considered dangerous, especially in the domain of games for health [Berkovsky et al., 2010]. To justify the need for investigating tailoring, the authors review persuasive strategies adopted in a range of existing domains of games for change: healthy eating [Kaipainen et al., 2012, Orji et al., 2013b], physical activity [Fujiki et al., 2008] and disease management [Brownson and Kumanyika, 2007]. By their own admission, this list is not exhaustive which gives great scope for future research in other domains. There is clearly little work on isolating the persuasive strategies employed in these games to study their effects on specific types of gamer which makes a strong case for their research. Further, from looking at these papers describing existing games, little to no explanation is given on the behavioural theories underpinning the choice of strategies.

[1] An annual festival exists for workshops and design classes on games for change: `www.gamesforchange.org`

To evaluate their research questions, Orji et al. implemented two versions of a custom model-driven online game called *Junk Food Aliens* (Figure 1) which targets the domain of healthy eating - players control an avatar and search for fruits and vegetables to save the planet from an invasion of junk foods. The reward-based version (JFA-R) adopted persuasive strategies such as achievement badges (Figure 3) whereas the competition-based version (JFA-C) adopted comparative strategies such as leaderboards displaying fictional scores of opponents (Figure 2).

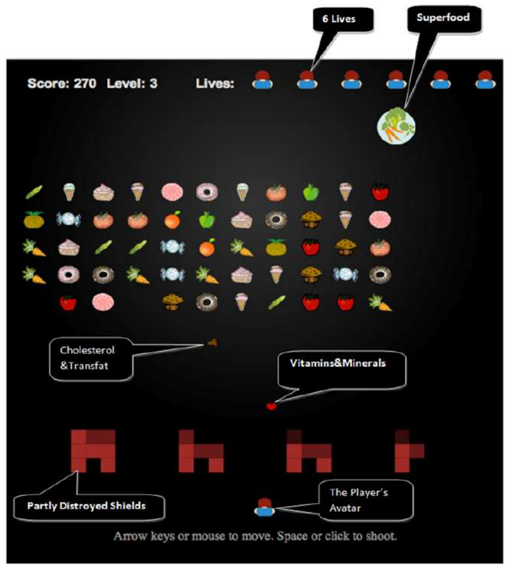
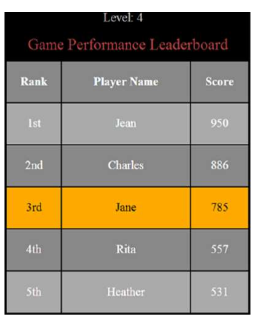


Figure 1: "Junk Food Aliens" (JFA): A persuasive game designed to change gamer behaviour towards healthy eating.


Level: 4

Game Performance Leaderboard

Rank	Player Name	Score
1st	Jean	950
2nd	Charles	886
3rd	Jane	785
4th	Rita	557
5th	Heather	531

Figure 2: JFA-C: Competition-based version of JFA.

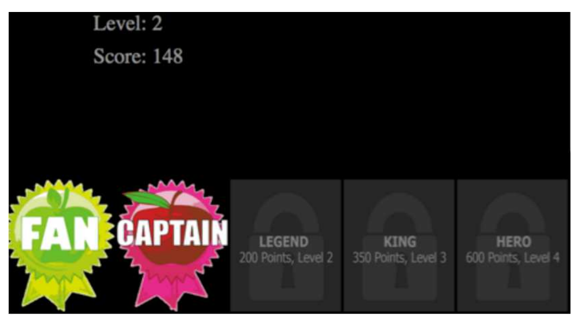


Figure 3: JFA-R: Reward-based version of JFA.

Main contributions of Orji et al. [2017] are fourfold. First, it is important to tailor games for change in order to observe any efficacy (the main experiment demonstrates this by measure of change in values of precursors of behavioural change - *attitude*, *self-efficacy*, and *intention*). Second, tailoring can be achieved by modifying the persuasive strategies adopted, rather than the mechanics of the game which minimises the costs involved in tailoring. Thirdly, the positive effect of tailoring was not mediated by an improved player experience (if not considered, this would be a potentially influential confounding factor). Finally, the adoption of a single persuasive strategy is sufficient to observe these effects - combining additional strategies is unpredictable and may be somewhat counter-productive.

Justifications for Conclusions

The authors' initial work suggested that an effective persuasive strategy for one type of gamer could be detrimental to another by means of a lower, or negative response to motivation levels [Orji et al., 2013a] (Table 1). This taxonomy forms a solid model to base the development of the JFA on, mapping independent gamer types defined by the BrainHex model [Nacke et al., 2014] against 8 common and pairwise distinct persuasive strategies selected by Gerling et al. [2014]. Results are based on a large-scale online survey of 1108 gamers who indicated how likely each of 10 selected strategies would influence eating decisions, using a validated scale [Drozd et al., 2012] before completing the BrainHex questionnaire to indicate their gamer type. The data collected from this initial experiment therefore forms as valid justification for the later comparison between Achievers (motivated by Reward) and Conquerors (motivated by Competition) as they are almost polar opposites in their responses to these strategies. Use of a custom model over a more general model such as Fogg's Behavioural Model (FBM) [Fogg, 2002] increases its applicability to persuasive *games*, instead of persuasive technology in general. This is at the cost of a need to re-analyse the chosen model for every new domain explored.

Table 1: β values confusion matrix: Strength of motivation of different players that result from different strategies. Positive β values indicate that gamers of this type are motivated by the corresponding given strategy. Negative β values indicate demotivation, whilst an empty value indicates neither motivation nor demotivation [Orji et al., 2013a].

Strategies / Gamer type	CMPT/ CMPR	COOP	CUST	PERS	PRAS	SEMT/ SUGG	SIML	REWD
Achiever	-	.15	-	-	-	.10	-	.10
Conqueror	.25	-	-	.12	-	.12	.14	-
Daredevil	-.10	-	-	-	-	-.14	.11	-
Mastermind	.12	-	.10	.12	-	.14	.12	-
Seeker	.10	-	.19	.11	.10	-	-	-
Socializer	.11	.17	-.12	-	-.12	-.13	-	-
Survivor	.17	-.20	-.13	-	-	.27	-	-.14
CMPT/CMPR = competition and comparison, COOP = cooperation, CUST = customization, PERS = personalization, PRAS = praise, SEMT/SUGG = self-monitoring and suggestion, SIML = simulation, REWD = reward.								

To empirically evaluate this model, for the rest of the paper the authors focus on games for healthy eating and restrict considerations on the effects two persuasive strategies (Competition and Reward) on two types of gamers (Achiever and Conqueror). Whilst somewhat limiting in scope, this approach allowed direct comparison between types of gamers shown empirically to be significantly-distinct in their response to these strategies [Orji et al., 2013a]. Furthermore, these strategies are common to existing games [Bell et al., 2006] and these player types are equally as dominant [Bartle, 1996].

Participants of the main experiment were recruited online via Amazon's Mechanical Turk (AMT) using the guidance of Mason and Suri [2012]. The authors recognise this trade-off between access to a large selection pool in return for greater difficulty ensuring participant motivation or attentiveness, leading to a potential loss of generalisability in results. After two pilot studies ($N = 40, N = 4$) verified experimental and instrumental validity, a large-scale randomized controlled online study of 272 valid participants (117 female, 155 male - 50% Achievers and 50% Conquerors, all game-players) investigated the effects of tailoring and contra-tailoring on these two types of gamer in a between-subjects design. In this sense, half of Achievers played JFA-R (tailored) and the remaining half played JFA-C (contra-tailored), and vice versa for the Conquerors (tailored played JFA-C and contra-tailored played JFA-R). Participant selection was weakened by its focus on self-selected experienced gamers who are not the prime target audience for games for change [Brox et al., 2011].

Before playing, participants completed a BrainHex survey [Nacke et al., 2014], demographic survey and responded to 3 scales to measure their baseline attitude, intention and self-efficacy towards healthy eating based on Ajzen [2002]. Once gameplay began, interruptions were displayed after each minute to show rewards or the leaderboard in JFA-R and JFA-C respectively. After participants finished playing their assigned version of JFA for 20 minutes, they responded to the same 3 scales to measure their post-task attitude, intention and self-efficacy towards healthy eating. Finally, they also assessed their experience (enjoyment, effort, competence and tension) [Ryan et al., 2006]. Results showed tailoring the game to a specific gamer type increased effectiveness (measured by positive changes in attitudes, intentions and self-efficacy towards healthy eating) while contra-tailoring did not (Figure 4).

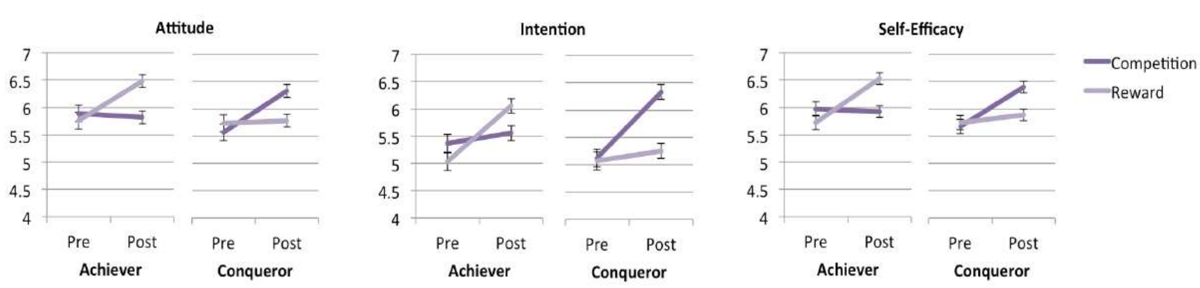


Figure 4: Mean values ± SE for Attitude, Intention, and Self-Efficacy by Gamer type (Achiever, Conqueror) and Game version (Competition, Reward).

Parallel mediated regressions [Hayes, 2013] showed the positive effects of tailoring were not mediated by improved player experience which provides strong support in favour of tailoring games to player type (Figure 5).

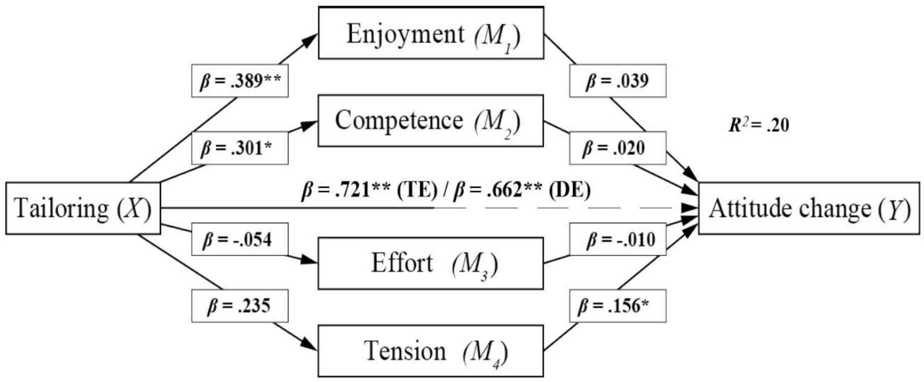


Figure 5: Parallel mediation model of tailoring on attitude change with play experience as mediator.

There is a respectable amount of evidence to support the authors' claims. The number and representativeness of the participant sample has been carefully considered, with the limitations of using AMT discussed.

Limitations and Suggested Further Work

Whilst the findings of the main experiment are convincing, well designed and validated with a good report on parametric data gathered, there are a series of limitations and unanswered questions that must be considered.

Orji et al. appreciate that scoping their main experiment to focus only on Achievers and Conquerors under Reward and Competition strategies respectively should encourage future research on different combinations of strategies and different gamer types. Their current research limits itself to the measure of change in attitude, intention, and self-efficacy toward healthy eating which immediately prompts questions about its effects on *tangible* changes in behaviour, rather than effects on *mediators* or *precursors* of behaviour change. Also, although purely illustrative, the domain choice of healthy eating should not limit investigations - by applying their model to other domains such as disease management and physical activity, it will be possible to decipher whether the effects of the persuasive strategy employed are domain-dependent. That said, the custom model cannot be immediately generalised to domains other than healthy eating without careful further investigation.

Ideally, participant recruitment should be intrinsically motivated rather than via AMT where not all participants may be psychologically engaged in tasks affecting generalisability of results. It would also be novel to explore the adoption of a single persuasive strategy compared with use of multiple in a given game, to observe whether compounding effects are positive or negative on efficacy of behaviour change. Now that we are aware that player experience does not mediate the observed benefits, it would be wise to also consider whether *game performance* as a potential mediating factor - players who perform better in the game under a given strategy may experience more positive feeling towards the tailored game which could influence the observed results. The current approach is also somewhat static in nature once participants are assigned to a given gamer type. Answers to the BrainHex questionnaire are what determines a player's type. Instead, it might be interesting to explore the avenue of determining a player's type in a more dynamic sense, perhaps during actual gameplay based on activity or game events. Future research should also consider effects on non-gamers seeing as many existing persuasive games cater for this group - e.g. the elderly population [de Oliveira et al., 2010].

Another novel suggestion to build on these findings is to experiment with existing gamification applications for physical exercise such as Apple Watch and Nike+ GPS, instead of focusing purely on games for change. By expanding into this area, this might require re-evaluation of the strategies originally selected by Gerling et al. [2014].

Conclusion

Orji et al. [2017] make a convincing case for personalising persuasive strategies employed in future games for change in order to observe effective change in behaviour. Whilst there are limitations to their work, it lays strong foundations for future work investigating tailoring games for change in other domains. Findings have potentially wider consequences in Persuasive Technology - with recommendations on how to tailor to specific groups of users whilst minimising the efforts and costs involved in doing so, designers of these technologies could better-understand the considerations needed when matching a persuasive strategy to their specific user group.

Word count: 1646 words (not inc. Citations, Figures or References)

References

Icek Ajzen. Constructing a tpb questionnaire : Conceptual and methodological considerations september , 2002. 2002.

Richard Bartle. Hearts, clubs, diamonds, spades: Players who suit muds. 06 1996.

Marek Bell, Matthew Chalmers, Louise Barkhuus, Malcolm Hall, Scott Sherwood, Paul Tennent, Barry Brown, Duncan Rowland, Steve Benford, Mauricio Capra, and Alastair Hampshire. Interweaving mobile games with everyday life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 417–426, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: 10.1145/1124772.1124835. URL `http://doi.acm.org/10.1145/1124772.1124835`.

Shlomo Berkovsky, Mac Coombe, Jill Freyne, Dipak Bhandari, and Nilufar Baghaei. Physical activity motivating games: Virtual rewards for real activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 243–252, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-929-9. doi: 10.1145/1753326.1753362. URL `http://doi.acm.org/10.1145/1753326.1753362`.

Ross C. Brownson and Shiriki Kumanyika. *Obesity Prevention: Charting a Course to a Healthier Future*, pages 515–528. Springer US, Boston, MA, 2007. ISBN 978-0-387-47860-9. doi: 10.1007/978-0-387-47860-9_22. URL `https://doi.org/10.1007/978-0-387-47860-9_22`.

Ellen Brox, Luis Fernandez-Luque, and Torunn Tøllefsen. Healthy gaming – video game design to promote health. *Applied clinical informatics*, 2:128–42, 04 2011. doi: 10.4338/ACI-2010-10-R-0060.

Marc Busch, Elke Mattheiss, Rita Orji, Andrzej Marczewski, Wolfgang Hochleitner, Michael Lankes, Lennart E. Nacke, and Manfred Tscheligi. Personalization in serious and persuasive games and gamified interactions. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, pages 811–816, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3466-2. doi: 10.1145/2793107.2810260. URL `http://doi.acm.org/10.1145/2793107.2810260`.

Rodrigo de Oliveira, Mauro Cherubini, and Nuria Oliver. Movipill: Improving medication compliance for elders using a mobile persuasive social game. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, UbiComp '10, pages 251–260, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-843-8. doi: 10.1145/1864349.1864371. URL `http://doi.acm.org/10.1145/1864349.1864371`.

Filip Drozd, Tuomas Lehto, and Harri Oinas-Kukkonen. Exploring perceived persuasiveness of a behavior change support system: A structural model. In Magnus Bang and Eva L. Ragnemalm, editors, *Persuasive Technology. Design for Health and Safety*, pages 157–168, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. ISBN 978-3-642-31037-9.

B. J. Fogg. Persuasive technology: Using computers to change what we think and do. *Ubiquity*, 2002 (December), December 2002. ISSN 1530-2180. doi: 10.1145/764008.763957. URL `http://doi.acm.org/10.1145/764008.763957`.

Yuichi Fujiki, Konstantinos Kazakos, Colin Puri, Pradeep Buddharaju, Ioannis Pavlidis, and James Levine. Neat-o-games: Blending physical activity and fun in the daily routine. *Comput. Entertain.*, 6(2):21:1–21:22, July 2008. ISSN 1544-3574. doi: 10.1145/1371216.1371224. URL `http://doi.acm.org/10.1145/1371216.1371224`.

Kathrin Maria Gerling, Regan L. Mandryk, Max Valentin Birk, Matthew Miller, and Rita Orji. The effects of embodied persuasive games on player attitudes toward people using wheelchairs. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 3413–3422, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2473-1. doi: 10.1145/2556288.2556962. URL `http://doi.acm.org/10.1145/2556288.2556962`.

Andrew F. Hayes. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. The Guilford Press, 01 2013. ISBN 9781609182304.

Kirsikka Kaipainen, Collin R Payne, and Brian Wansink. Mindless eating challenge: Retention, weight outcomes, and barriers for changes in a public web-based healthy eating and weight loss program. *Journal of medical Internet research*, 14:e168, 11 2012. doi: 10.2196/jmir.2218.

Rilla Khaled, Ronald Fischer, James Noble, and Robert Biddle. A qualitative study of culture and persuasion in a smoking cessation game. In Harri Oinas-Kukkonen, Per Hasle, Marja Harjumaa, Katarina Segerståhl, and Peter Øhrstrøm, editors, *Persuasive Technology*, pages 224–236, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. ISBN 978-3-540-68504-3.

Winter Mason and Siddharth Suri. Conducting behavioral research on amazon's mechanical turk. *Behavior Research Methods*, 44(1):1–23, Mar 2012. ISSN 1554-3528. doi: 10.3758/s13428-011-0124-6. URL `https://doi.org/10.3758/s13428-011-0124-6`.

Lennart E. Nacke, Chris Bateman, and Regan L. Mandryk. Brainhex: A neurobiological gamer typology survey. *Entertainment Computing*, 5(1):55 – 62, 01 2014. ISSN 1875-9521. doi: https://doi.org/10.1016/j.entcom.2013.06.002. URL `http://www.sciencedirect.com/science/article/pii/S1875952113000086`.

Rita Orji, Regan L. Mandryk, Julita Vassileva, and Kathrin M. Gerling. Tailoring persuasive health games to gamer type. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2467–2476, New York, NY, USA, 2013a. ACM. ISBN 978-1-4503-1899-0. doi: 10.1145/2470654.2481341. URL `http://doi.acm.org/10.1145/2470654.2481341`.

Rita Orji, Julita Vassileva, and Regan L. Mandryk. Lunchtime: a slow-casual game for long-term dietary behavior change. *Personal and Ubiquitous Computing*, 17(6):1211–1221, Aug 2013b. ISSN 1617-4917. doi: 10.1007/s00779-012-0590-6. URL `https://doi.org/10.1007/s00779-012-0590-6`.

Rita Orji, Regan L. Mandryk, and Julita Vassileva. Improving the efficacy of games for change using personalization models. *ACM Trans. Comput.-Hum. Interact.*, 24(5):32:1–32:22, October 2017. ISSN 1073-0516. doi: 10.1145/3119929. URL `http://doi.acm.org/10.1145/3119929`.

Wei Peng. Design and evaluation of a computer game to promote a healthy diet for young adults. *Health Communication*, 24(2):115–127, 2009. doi: 10.1080/10410230802676490. URL `https://doi.org/10.1080/10410230802676490`. PMID: 19280455.

Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, 30(4):344–360, Dec 2006. ISSN 1573-6644. doi: 10.1007/s11031-006-9051-8. URL `https://doi.org/10.1007/s11031-006-9051-8`.

Critical Review: Exploring Interactions with Physically Dynamic Bar Charts

Introduction

Studies investigating how data can be effectively presented to, explored and interpreted by users forms the core part of Information Visualisation ('InfoVis') to support users in the decision-making process [Card and Mackinlay, 1997]. This review summarises and critically analyses Taher et al. [2015] whose paper explores the use of physically dynamic bar charts as devices for exploring user interactions with visualisations of data, to determine future work in this domain of Information Visualisation.

Summary of Contributions

Taher et al. extend existing work on physical visualisations (*physicalizations*) [Jansen et al., 2015] to investigate how users interact with *physically dynamic* bar charts as a way of exploring and manipulating shape-changing visualisations of datasets. Existing work into physicalizations involve problematic *static* models that do not respond to user interactions [Jansen et al., 2013] and are therefore "*disconnected*" from the data source when created. With the advent of shape-changing technology [Rasmussen et al., 2012], there is scope for the manufacture of physically dynamic displays to help decision makers reason about and manipulate data sets in a non-virtual and non-static way. This motivation leads Taher et al. to explore the ways users interact with data displayed in this mode to understand whether *physical* interactions (such as touching specific bars) or *gestures* (such as swiping a touch-screen) or a *combination* of the two is more intuitive to users when solving common problems. Whilst the authors concede that physical dynamic visualisations are not new [Leithinger and Ishii, 2010, Follmer et al., 2013], they claim there is little analysis into effective *interactions* with data of *dynamic* physical modality, unlike the abundance of work investigating their static counterparts [Stusak and Aslan, 2014].

The point system to support the author's research is *EMERGE* - a 10×10 physical bar chart which uses a set of dynamic self-actuating rods with an RGB display projected onto it (Figure 1). An immediate and obvious limitation of using a Bar Chart point system is that any conclusions drawn about its effectiveness in physically-dynamic form cannot be generalised to other types of InfoVis systems, such as Dynamic Histograms, Parallel Coordinates and Theme Rivers without further research.

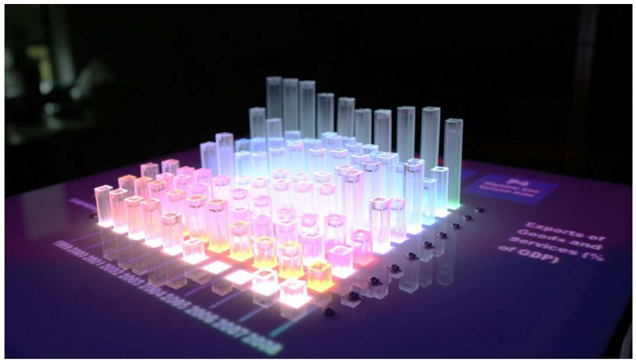


Figure 1: EMERGE: Exploring Interactions with Physically Dynamic Bar Charts using actuating physical rods and RGB LEDs to display international export data.

EMERGE allows users to interact with the dataset it represents using a set of 4 interaction-sets derived from subcategories of the taxonomy of interactive dynamics for visual analysis described by Heer and Shneiderman [2012] - *annotation*, *filtering*, *organisation* and *navigation* (Table 1). Heer and Shneiderman lay out 3 high-level categories in their model - *Data and View Specification*, *View Manipulation*, and *Process and Provenance*. (Figure 2). In this sense, whilst the selection of InfoVis model is careful and grounded in background theory, the choice of subcategories by Taher et al. for interacting with EMERGE is somewhat arbitrary and limited in scope with little justification, which immediately invites further research into different forms of interactions with physicalisations from the taxonomy.

*Table 1: Task-sets and interaction techniques explored during the user study with EMERGE: annotation, filtering, organisation and navigation with the category of Heer and Shneiderman [2012] in **bold**.*

Task	**Overview**	**Interaction Techniques**
Annotation ***(Process & provenance)***	Selecting and marking individual data points.	Point, pull, press.
Filtering ***(Data view & specification)***	Hiding and refining data for enhanced perception and comparison.	Swipe away, manual press, assisted press, press shortcut, and press to compare.
Organization ***(View manipulation)***	Data arrangement by moving rows and columns.	Drag and drop with immediate transition and hide-all with transition, press with instant transition and hide-all with transition.
Navigation ***(View manipulation)***	Controlling the view of large data sets.	Scroll, directional arrows, directional press, and paging.

Data and View Specification	**Visualize** data by choosing visual encodings.
	Filter out data to focus on relevant items.
	Sort items to expose patterns.
	Derive values or models from source data.
View Manipulation	**Select** items to highlight, filter, or manipulate them.
	Navigate to examine high-level patterns and low-level detail.
	Coordinate views for linked, multidimensional exploration.
	Organize multiple windows and workspaces.
Process and Provenance	**Record** analysis histories for revisitation, review, and sharing.
	Annotate patterns to document findings.
	Share views and annotations to enable collaboration.
	Guide users through analysis tasks or stories.

Figure 2: Taxonomy of interactive dynamics described by Heer and Shneiderman [2012].

Main contributions of Taher et al. [2015] are threefold. First, a set of 14 potential interactions (both physical and gesture based) for manipulating and exploring data presented in dynamic physical bar charts such as EMERGE. Second, the findings of their user study ($N = 17$) evaluates which of the interactions are effective and intuitive in completing a set of analysis tasks, and which interactions match users' initial preconceptions for how to achieve these tasks. Finally, a set of design considerations to advise future research on the challenges of presenting data in physically dynamic form. Overall, we learn that a *combination* of gestures and physical interactions are effective. Smaller interactions such as annotation of specific data points can be afforded by physical interaction whereas larger interactions such as organization can be afforded touch-screen gestures.

Justifications for Conclusions

Taher et al. set about their investigations by creating 14 *baseline* interactions across the 4 main tasks (annotation, filtering, organisation, navigation - e.g. Figure 3 and 4) for users to interact with EMERGE. The authors avoided early experimentation to generate different types of interaction for each task before the main study as existing research forewarned against this due to the immaturity of the area [Hornbæk, 2013]. Whilst this was sensible to consider, the mechanics of the baseline interactions used in the study are strongly-coupled with the hardware capabilities of EMERGE and no explanation is given by the authors into where the inspiration for each proposed interaction came from, which causes some early concern over generalisability of results to other implementations of physically dynamic bar charts.

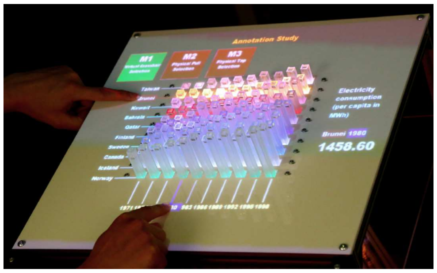


Figure 3: Annotation (Point technique).

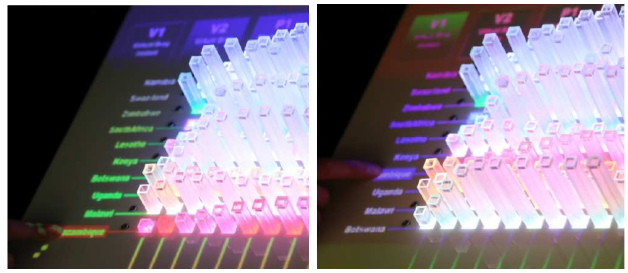


Figure 4: Organisation (Drag and Drop technique).

A within-subjects study (N = 17, 6 female, mean age = 27) investigated the authors' research. Participants had no experience with shape changing displays but no detail is given about their selection criteria or origin (i.e. the selection pool) which means it is difficult to comment on results being representative. The small sample size, lack of representative sampling with respect to gender and use of extrinsic (monetary) motivation are all limiting factors on the study's external validity. After participants were individually introduced to EMERGE, they were asked how they initially thought they could achieve each task. Then, each technique for a given task was demonstrated before a verbal data analysis task was set for each interaction (e.g. "*Select the year and country with the highest electricity consumption.*") to observe user behaviours and feedback.

Counter-balancing on the ordering of task-sets and techniques reduced order bias, but it could be argued that the within-subjects design caused a training effect as users became more comfortable with the EMERGE interface. After each task, a 5-point Likert questionnaire was carried out (Figure 5) along with elicitation of users' comments. In this sense, the study was well-designed, allowing for initial preconceptions of interactions to be captured before any demonstrations took place. That said, different datasets were used for each task, which could have varying effects on the participants' ability to interpret the visualisation and therefore influence their rating of the interaction under scrutiny. Also, besides the Likert ratings the data collected is purely qualitative and so more parametric data such as performance metrics (task completion times) would strengthen the authors' conclusions.

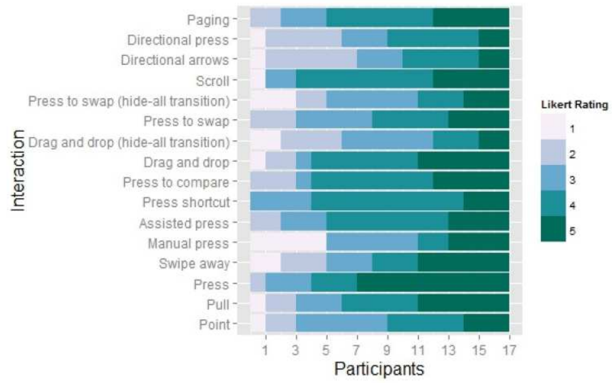


Figure 5: Likert scale ratings for helpfulness of interaction techniques. Range = 1: Strongly Disagree, 5: Strongly Agree.

EMERGE's implementation presents a large limitation in the authors' findings. Firstly, the chart itself is of a 3D modality with no vertical axis. The only meaningful vertical information is ascertained by visual comparison of bar heights rather than reading of direct data. Secondly, the representation of larger datasets would require a far more data points than the 10×10 grid currently allows for - it is likely that for higher resolution of data points, navigation interactions such as pagination would suffice whereas pressing data points to navigate between single rows would be cumbersome. It is also likely that new background theory may need to be considered with higher resolution of data points - perhaps use of the *overview first, zoom and filter, then details on demand* mantra described by Shneiderman [1997]. Thirdly, specific technical challenges discussed (actuation speed and noise, rod spacing, size of setup) may have also mediated or impacted the results in ways that were unaccounted for. For example, almost all participants reacted hesitantly due to these speed and noise of actuation. Further analysis (e.g. use of parallel mediated regression testing) should investigate this.

Limitations and Suggested Further Work

Taher et al. [2015] present a respectable investigation of potential techniques to interact with data presented in a physically dynamic form and their conclusion on the effectiveness of adopting *both* physical and gesture-based interactions with these systems seems valid, but a set of limitations with their work leads to some questions being open to future research.

Crucially, the number of participants involved should be increased and to ensure conclusions are generalisable with external validity, the sample should be representative of the general population (approx. 50% gender split). Also, aside from the Likert scale data, the paper lacks a lot of parametric data which might provide further insight into efficacy of interactions - further controlled studies could attempt to measure performance metrics such as task completion times and accuracy for each interaction type.

Seeing as the investigation focuses on a narrow subset of the interaction taxonomy of Heer and Shneiderman [2012] on a Bar Chart, the obvious extensions to the authors' work would be to explore different interaction types such as Record, Co-ordinate and Share using EMERGE, and extend experiments across different point systems such as Dynamic Histograms, Parallel Coordinates and Theme Rivers. As *both* physical and gesture based interactions are effective for small and large grained tasks respectively, it would also be wise to see how *combining* both interactions into one activity affects findings.

The InfoVis design space is very large [Card and Mackinlay, 1997], so it would be novel to explore how users perceive dynamic physicalisations which represent complex datasets which *change* in real time - e.g. social network data [Federico et al., 2011]. As well as responding to human interaction via shape-changing interfaces, these datasets could also change of their own accord which opens up more questions into how users would respond to this type of visualisation. Another novel suggestion would be to apply these visualisations to a specific context or domain, to compare efficacy of interactions across modalities (i.e. *static physical* versus *dynamic physical* versus *virtual*). An example might be in an educational setting where kinaesthetic techniques are encouraged [Pourhosein Gilakjani, 2011]. Furthermore, investigations could also be extended to observe the effects of using dynamic physicalisations whilst a user is immersed in a VR environment, to combat a lack of 'presence' in these settings and afford data manipulation in virtual settings with physical assistance [Tennent et al., 2017]. Finally, existing work analysing use of static physicalisations for the blind [Perkins, 2002] can be extended to understand how dynamic physicalisations can offer further support in data manipulation by the visually impaired.

Conclusion

Taher et al. [2015] set the foundations for future investigations into use of shape-changing displays for accomplishment of common InfoVis tasks. Their research is limited in several ways by the implementation of the EMERGE system, but raises important design considerations for future work investigating dynamic physicalisations in other domains. Findings have potentially wider consequences in the topic of Information Visualisation - with recommendations on how to design systems supporting the fundamental interactions such as those described by Heer and Shneiderman [2012], these visualisations can serve as effective data analysis tools in a variety of domains.

Word count: 1598 words (not inc. Citations, Figures or References)

References

S. K. Card and J. Mackinlay. The structure of the information visualization design space. In *Proceedings of the 1997 IEEE Symposium on Information Visualization (InfoVis '97)*, INFOVIS '97, pages 92–, Washington, DC, USA, 1997. IEEE Computer Society. ISBN 0-8186-8189-6. URL `http://dl.acm.org/citation.cfm?id=857188.857632`.

Paolo Federico, Wolfgang Aigner, Silvia Miksch, Florian Windhager, and Lukas Zenk. A visual analytics approach to dynamic social networks. 09 2011. doi: 10.1145/2024288.2024344.

Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. inform: Dynamic physical affordances and constraints through shape and object actuation. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pages 417–426, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2268-3. doi: 10.1145/2501988.2502032. URL `http://doi.acm.org/10.1145/2501988.2502032`.

Jeffrey Heer and Ben Shneiderman. Interactive dynamics for visual analysis. *Commun. ACM*, 55(4):45–54, April 2012. ISSN 0001-0782. doi: 10.1145/2133806.2133821. URL `http://doi.acm.org/10.1145/2133806.2133821`.

Kasper Hornbæk. Some whys and hows of experiments in human–computer interaction. *Found. Trends Hum.-Comput. Interact.*, 5(4):299–373, June 2013. ISSN 1551-3955. doi: 10.1561/1100000043. URL `http://dx.doi.org/10.1561/1100000043`.

Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. Evaluating the efficiency of physical visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2593–2602, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1899-0. doi: 10.1145/2470654.2481359. URL `http://doi.acm.org/10.1145/2470654.2481359`.

Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. Opportunities and challenges for data physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3227–3236, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3145-6. doi: 10.1145/2702123.2702180. URL `http://doi.acm.org/10.1145/2702123.2702180`.

Daniel Leithinger and Hiroshi Ishii. Relief: A scalable actuated shape display. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '10, pages 221–222, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-841-4. doi: 10.1145/1709886.1709928. URL `http://doi.acm.org/10.1145/1709886.1709928`.

Chris Perkins. Cartography: progress in tactile mapping. *Progress in Human Geography*, 26(4):521–530, 2002. doi: 10.1191/0309132502ph383pr. URL `https://doi.org/10.1191/0309132502ph383pr`.

Abbas Pourhosein Gilakjani. Visual, auditory, kinaesthetic learning styles and their impacts on english language teaching. *Journal of Studies in Education*, 2:104, 12 2011. doi: 10.5296/jse.v2i1.1007.

Majken K. Rasmussen, Esben W. Pedersen, Marianne G. Petersen, and Kasper Hornbæk. Shape-changing interfaces: A review of the design space and open research questions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 735–744, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi: 10.1145/2207676.2207781. URL `http://doi.acm.org/10.1145/2207676.2207781`.

Ben Shneiderman. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 3rd edition, 1997. ISBN 0201694972.

Simon Stusak and Ayfer Aslan. Beyond physical bar charts: An exploration of designing physical visualizations. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, pages 1381–1386, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2474-8. doi: 10.1145/2559206.2581311. URL `http://doi.acm.org/10.1145/2559206.2581311`.

Faisal Taher, John Hardy, Abhijit Karnik, Christian Weichel, Yvonne Jansen, Kasper Hornbæk, and Jason Alexander. Exploring interactions with physically dynamic bar charts. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3237–3246, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3145-6. doi: 10.1145/2702123.2702604. URL `http://doi.acm.org/10.1145/2702123.2702604`.

Paul Tennent, Joe Marshall, Brendan Walker, Patrick Brundell, and Steve Benford. The challenges of visual-kinaesthetic experience. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, DIS '17, pages 1265–1276, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4922-2. doi: 10.1145/3064663.3064763. URL `http://doi.acm.org/10.1145/3064663.3064763`.