Audialization in Serious Games Analytics: Visualizing Player Performance Improvement by Sound or Music

Christian Sebastian Loh
Virtual Environment Lab (V-Lab)
625 Wham Drive, Mailcode 4610
Southern Illinois University
Carbondale, IL 62901-4610
csloh@siu.edu

Aaron Ekstrand
Virtual Environment Lab (V-Lab)
625 Wham Drive, Mailcode 4610
Southern Illinois University
Carbondale, IL 62901-4610
lintamacar@gmail.com

ABSTRACT
A well-designed game can be used to modify players' actions and behaviors. In performance improvement, learners are often required to practice a task over and over again until they become proficient at it (i.e., improved competency) – a process that is highly similar to 'grinding' in digital games. Competency can be directly measured as the change in player's course of action or behaviors, using similarity measures. In most analytics, graphical representation of the information (i.e., visualization) is often the only way to communicate insights to stakeholders. Although auditory representation can be just as important in learning, it is often under-utilized in visualization of information. We introduce 'audialization' – i.e., audial visualization, to visualize the improvement of players' behaviors and competency by sound or music. Such multimodal representation can be useful in many ways, including game design, human-computer interaction, and Serious Games analytics.

Author Keywords
Serious Games analytics; audialization; similarity measures; expert-novice performance; competency improvement.

ACM Classification Keywords
D.4.8 [Performance]: Measurements; H.5.2 [User Interfaces]: Auditory (non-speech) feedback; I.5.3 [Clustering]: Similarity Measures

INTRODUCTION
Well-designed games have been shown to modify players' actions and behaviors. In human performance literature, learning is often associated with changes in behavior [1]. In fact, the process of 'grinding' (to level up) in digital games is highly similar to the process of performance improvement in training because both processes require a person to perform a task over and over again until s/he becomes proficient. In the case of a gamer, grinding results in an increase in level, and in the case of a learner, practice results in competency improvement that is measurable as changed course of actions. It should be noted that 'grinding' is not suitable for serious games. Even though it involves repeated performance, there is no guarantee that the flat points awarded per task (as in digital games) is commensurate with training performance in real-life. In serious games, a better methodology is needed to measure competency and assess the training performance of player-trainees – competencies validated against real experts (instead of a flat point system).

While 'best time' often represents 'best performance' in competitive (sports and) games, striving to achieve best time in games that are designed for learning and behavior modification can often lead to counter-productive or detrimental outcomes. People working under time pressure have been shown to be tempted into making hasty decisions or taking chances [2]–[4]. Once these 'risky behaviors' become inculcated, they can easily result in poor decision habits and workplace disasters, if left unchecked [5]. Games created as tools for empowerment and policy improvement – e.g., disaster preparation, surgery, learning to driving, job interview preparation, migrant familiarizing with new environment – may benefit from a competency-based approach, instead of an over-reliance on the 'best time' criterion.

MOTIVATION

Expert-Novice Behavioral and Performance Differences
Findings in the area of expertise found experts to possess different reasoning patterns, decision-making procedures, and significantly better problem-solving strategies than novices [6]. People's belief systems, which affect their actions, have also been found to be differ between experts and novices [7].

Likewise, literature in training psychology [8], [9] revealed novices to exhibit a tendency to follow rules blindly (when solving problems) because they have yet to acquire the context in which those rules operate. As they gradually learn

DOI:

to apply the right rules with the right conditions, they are said to be growing in their *competency*.

Similarity Measure

Similarity measure was originally developed in Record Linkage analysis [10] to statistically determine if two data sets might be duplicates [11]. Since then, the method has been incorporated into many different fields of research ranging from facial detection, to traffic analysis, to genetic sequencing.

Until recently [12], [13], similarity measures have never been used to determine performance differences in human behaviors. Readers are referred elsewhere (e.g., Wikipedia) for more details on Similarity Measures as it is beyond the scope of the paper. We will instead, limit our discussion in this study to the Jaccard coefficient as an example. Jaccard coefficient of two sample sets can be determined by dividing the size of their intersection by the size of their union:

$$Jaccard\ (A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

The resulting value for two completely different sets is 0, whereas the value for identical sets is 1. This allows them to be easily transformed into 0% and 100%, the values of which are highly convenient for communication with stakeholders and nonprofessionals.

Measuring Performance Improvement in Serious Games

In performance improvement research, people's competency in a particular task can be understood as a demonstrable and measureable change in their 'course of actions' [8], [9]. In Serious Games, such changes can be easily tracked using telemetry, while the (dis)similarities between the action-sequences calculated as similarity measures to discriminate experts from novices [12], [13].

For example, consider a case involving a known expert and a player whose performance is yet to be determined. If the similarity measure (using the Jaccard coefficient) between the expert and the player is 1, the player is a 'likely-expert', whereas 0 classifies the player as a novice, due to his/her actions being completely dissimilar to that of the expert. If the similarity measure (using the Jaccard coefficient) between the expert and the player is 1, the player is a 'likely-expert', whereas 0 classifies the player as a novice, due to his/her actions being completely dissimilar to that of the expert. One should bear in mind that players' true level of performance should always be confirmed using multiple trials, and not with a one-shot experiment. (For cases involving multiple experts in one scenario, see [14].)

SERIOUS GAMES RESEARCH METHODS

The Good: Repeated Measure Studies

Competency-based measurement has a wide application in Serious Games analytics because it can be used for the comparison of performances from two different instances (e.g., round 1 vs. round 2), instead of just comparing between classes of players (i.e., expert(s) vs. novices). This is known in quantitative research as a repeated-measure study (RMS) method, where players act as their own (internal) controls. The RMS is especially important for serious games (effectiveness) research because it moves researchers away from the frequently reported, but flawed, media comparison study (MCS) method, prevalent in many fields of research involving learning with technology.

The Bad: Media Comparison Studies

The MCS method of research is one that involves the comparison between a group of (new) technology-users (say, serious games) and a control group (e.g., students from a traditional classroom). Such an approach is flawed because it essentially compares apples to oranges, so to speak. See [15]–[17] for treatises explaining why the MCS approach is a flawed and meaningless method to prove the effectiveness of technology in learning.

LEVELS OF COMPARISON

User-generated data (e.g., players' actions and behaviors) can be traced and collected *in situ* [18] using simple telemetry [19], [20], or more complex frameworks – such as *Information Trails* [21]–[23], which combines telemetry, data mining, and real-time visualization for Serious Games analytics.

The use of similarity measures to analyze user behaviors is not new and has been used for Web analysis [24], and game analytics [25]. It has recently been used to compare performance differences between experts and novices based on players' competency improvement in [12]–[14]. This approach is meaningful on several levels because it can be adapted for use in performance measurement, improvement, and assessment for serious games.

1. Instance Comparison: Player against Self

The comparison of two instances can be used in self-improvement because it reveals (to players) how one's course of actions changes over time/iterations of practice. This type of comparison between two gameplay instances was first seen in car-racing games in the form of 'ghost/ shadow driver,' which was a projected gameplay record from the previous round(s), to allow players to compete against their own 'shadow' for self-improvement.

2. Instance Comparison: Player against Player

Since it is possible to calculate all players' performance against the experts' baseline as a similarity index using similarity measures, the Expertise Index can be used to rank all players accordingly from 0 (novice) to 1 (likely-expert) [14], in the manner of a Leaderboard.

The rankings would allowed the players to identify where they stand in the ecosystem and, additionally, compare their performance against one another, as well as that of the expert(s). If the gameplay records are made public, trainees of lower ranking may be able to emulate the expert's performance (using a ghost/shadow option mentioned above) for training and learning, a scenario that is unheard of in today's training environment.

In addition, the ranked outcomes (based on competency levels of the player-trainees) can be visualized as a report to better communicate the effectiveness of the serious games to the stakeholders and serve as insights for serious games analytics. (See [12]–[14] for a lengthy explanation of the whole process.)

3. Instance Comparison: Team against Team

Last but not least, the similarity comparisons can also occur at the team level by transforming the combined indices of a team and comparing it against that of the opposing team (e.g., team competition and 'capture the flag').

VISUALIZATION OF INSIGHTS IN ANALYTICS

However, the values of these performance measurement approaches are highly dependent on information visualization – namely, the effective and efficient representation and communication of insights to stakeholders. At this juncture, the *de facto* method to visualize insights for most analytics – including (serious) games, is by way of graphical representation. An over-reliance on visual representations is understandable because the commercial world and the business (intelligence) industry favor the use of graphs to represent growth and earnings.

Visualizing Insights Effectively

Although there is nothing wrong with representing information visually, data analysts will soon run into design problems when more insights become available through the discovery of better analytics metrics and methods. Deciding which insights should be included on dashboards with limited screen-sizes (especially in the case of mobile devices) will likely be the focus of usability and user experience research for several years to come.

Visualizing Insights Efficiently

Instead of trying to increase the screen-size of devices (e.g., iPhone 6) to accommodate the growing amount of information and insights, one approach is to find more effective ways to design the Dashboard. Another equally viable alternative is to represent the insights more efficiently using other communication channels, ergo, sound.

AUDIALIZATION OF INSIGHTS

Human communications often involve both auditory and visual stimuli. Instead of representing the insights (from analytics) visually, an equally viable approach is to represent them using sound. Although human-computer interface research has many names for turning data into sound, such as auditory transformation, audification, and sonification [26]–[28], we felt that none of the terms depict what we want to say in terms of the representation of insights from analytics by directly affecting the message receiver audially, as in 'audial-visualization;' hence, the term: *audialization*.

As [29] pointed out, "auditory interfaces have the potential of making devices which rely solely on visual displays more usable and accessible to a wide range of users" (p.18). There are many ways and algorithms to convert data and information into sounds (e.g., MIDI, text-to-speech). Conceivably, insights from analytics can likewise be audialized into speech, monoaural tones, musical notes, or some combinations thereof (see [28] for a list of auditory transformation methods).

Message Representation with Sound or Music

As the usage of monoaural tones to indicate danger (e.g., police-car siren, ambulance), or discovery (e.g., metal detector, Geiger counter) are quite common in daily lives, it may also be more intuitive for understanding by stakeholders and nonprofessionals. For instance, [30] chose a tonal approach to represent the differences in the winning time of Olympic athletes in several programs.)

Using a tonal approach to representing information seemed reasonable and intuitive [26], [31]. However, instead of representing data one at a time (where a single tone would be adequate [30]), we needed to represent two sets of 'player's data' for immediate comparison – whether in tandem, or sequentially. Another challenge is that the choice of tone used in the audialization process must somehow convey the insights to stakeholder – i.e., presenting them with a sense of the 'distance' separating the performance between the two players, and must further do so throughout all levels of representation: (1) player vs. self, (2) player vs. player, and (3) team vs. team. To achieve this effect, we thought musical tones – with naturally existing chromatic distances, would better represent the audialization of Jaccard coefficients than monoaural tones.

In the following sections, we will describe our method – starting from the calculation of similarity index with action sequences for competency improvement, followed by the conversion of the similarity index into musical tone for audialization as Serious Games analytics.

METHODOLOGY

Materials and Participants

The serious game used in this study is a military 'ground reconnaissance' game with a *search and retrieve* mission comprised of seven checkpoints. This was a nonlinear game where a player was allowed to visit the checkpoints in any order. Upon entering the game, a mission-giver would brief the player with the objective to *search and retrieve* six villagers from a wooded area. The sequence by which the player visited the checkpoints were recorded as the 'action sequence' chosen for that round.

Upon completion of the initial *search and retrieve* round, players were asked to attempt a second *plan and execute* round. They were shown the map of the game with the locations of all checkpoints revealed. They were to make use of this new information to plan their route and make their checkpoint visitations more efficient. A player's competency for Round 2 (after planning) should improve considerably because they were given a chance to study the area map and (possibly) memorize the locations of the checkpoints.

We traced a total of 534,837 raw (gameplay) data points *in situ* using the *Information Trails* framework. A total of 62 players (55 novices, 7 experts) from a mid-western public university participated in this study. All players attempted the Challenge Round, but only 56 of them (49 novices, 7 experts) completed Round 2. (Six players, all of them novices, had to be dropped from Challenge Round due to technical issues and/or network problems.) We then calculated the Jaccard coefficients for all players against that of a particular expert. We fit the expert novice model using Partial Least Squares Discriminant Analysis and predicted the performance of the experts and novices in our data for both Round 1 and Round 2 [32]. We identified 17 novices who successfully 'crossed-over' into the expert group in Round 2. This is demonstrated in the improved competency for Round 2 as a whole, as seen in the higher median and mean in the Jaccard coefficients (Table 1).

	n	***Median***	***Mean***	***SD***
Round 1	62	.167	.321	.308
Round 2	56	.4	.533	.366

Table1: Distribution of Players' Jaccard Coefficients

Sonification

Sonication is the transformation process to convert data into sound [28]. Using the instructions provided by [33], we created a Python program [harmony.py] for the sonification process to transform the Jaccard coefficients into a set of sine waves. Using the following equation for oscillation:

$$y(t) = A \times \sin(2\pi f t + \varphi) \qquad (2)$$

where A is the amplitude of the oscillation, f is its frequency; t is the length of time, and φ is the phase of the oscillation at $t = 0$.

We set a Jaccard coefficient of 0 to be equal to the frequency of 440Hz (note A4) [jaccard0.wav], and a Jaccard coefficient of 1 to the frequency of 880Hz (note A5) [Jaccard1.wav]. This process artificially fixed the total distance between the minimum and the maximum of Jaccard coefficient (0 to 1) as 440Hz – spanning the complete octave of A (A4 to A5). The frequency of a Jaccard coefficient of 0.5 can be calculated as: 440Hz + 440Hz × 0.5 = 660Hz.

Figure 1 shows the differences in frequency between the two sonified Jaccard coefficients, where higher Jaccard values will result in more oscillations per second (i.e., higher frequency and pitch).

AUDIALIZATION OF PLAYER PERFORMANCE

To effectively compare the Jaccard coefficients from two player instances, we would play the single notes (one for each player) sequentially as a 'double-tone.' For a team that comprised of more than one player, the sound file produce was comprised of *n* sine waves, where *n* is the number of players. The two waves would then be sounded sequentially to audialize the performance differences between two teams.

(For the following discussion, please refer to the audios embedded in *audialization.mov*.)

1. Level of Comparison: Player vs. Self

For example, a player who scored 0 in the first round and 0.2 in the second round was said to have a performance increment of 20%. Her changes in performance would be audialized using the double notes: 440Hz followed by 528Hz.

2. Level of Comparison: Player vs. Player

For example, a player who scored 0.5 against an expert would be audialized using the double notes: 660Hz followed by 880Hz.

3. Level of Comparison: Team vs. Team

We will illustrate this using the data from the study. We included all 62 players from Round 1 as Team 1, and all 56 players who completed Round 2 as Team 2.

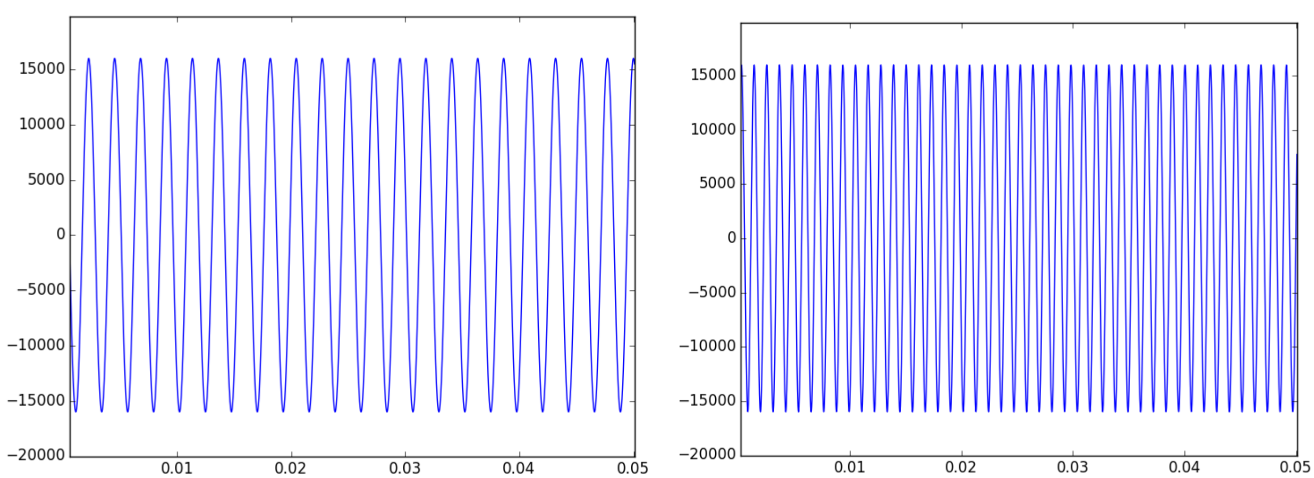

Figure 1. Sonified Jaccard coefficients: 0 (left) and 1 (right).

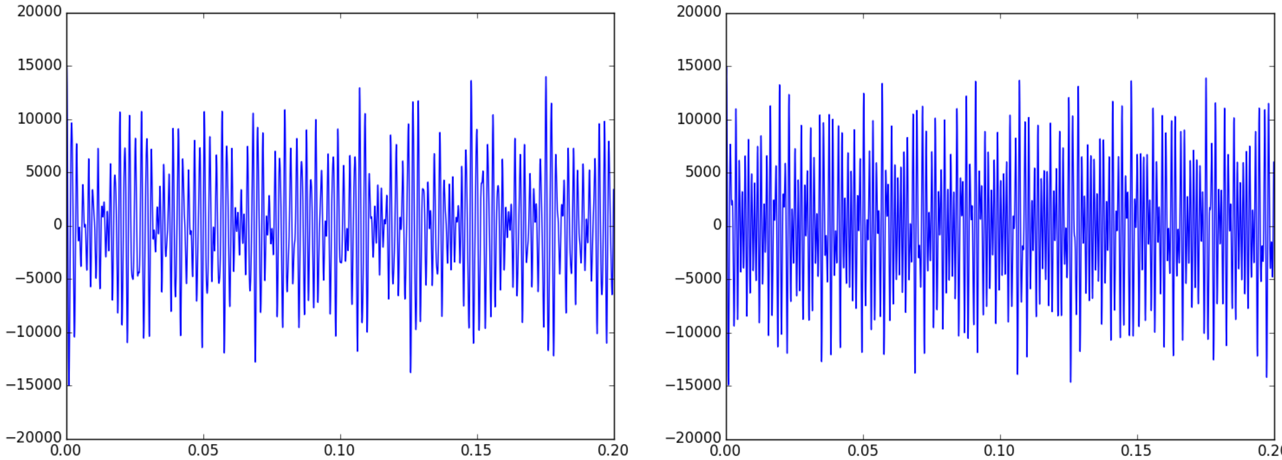

Figure 2. Team 1 (left) and Team 2 (right).

We use harmony.py to sonify the data and produced two files, which were then played one after another (like the above double-tone example) to audialize the findings.

Figure 2 shows the two sound waves: the sound wave from Team 2 has higher Jaccard values, as indicated by more oscillations per second, and higher frequency and pitch.

Additional Audialization Parameters

Spatial separation is another means by which people 'visualizes' sound. For instance, the musical notes mentioned in the last segment can be audialized through channel separations to give its listeners a greater sense of 'distance between data.' One can easily perform the channel separation by panning the audial file for Team 1 to 'hard left', and the audial file for Team 2 to 'hard right,' to achieve a maximum (physical) separation of 180^o. [It should be noted that an audio separation of 'left to right' is much easier to accomplish then an audio separation of 'top to bottom.']

CONCLUSIONS

In the past, research involving audification and sonification processes was mostly conducted to better the lives of the visually impaired [31]. Visualization of insights in the Big Data/Analytics is a necessary step forward because there are so much information produced that it is becoming almost impossible for human to 'see' the relationships within these data.

As game designers and researchers take advantage of player-generated data and convert them into game analytics for the improvement of gameplay experience [32, 33] and monetization, better means of information representation will be needed to communicate the insights to stakeholders and nonprofessionals. Serious games researchers have likewise begun looking for ways to convert player-generated data into Serious Games analytics for the purposes of performance measurement, assessment, and improvement [36].

As more data becomes available and are converted into insights to improve profit margins [37], user experience [32, 33], skills and performance [36], and for other social/humanitarian functions, the resulting amount of insight is sure to exceed the capacity of the visual Dashboards.

Given the amount of data collected and analyzable in today's world, an over-crowded dashboard can be counter-productive and difficult to understand because the visual information is competing for the same transmission channel. Multiple Resource theory [38] informs us that two tasks that use the same input channels (e.g. visual-visual) will interfere with each other more than tasks that use different channels (e.g., visual-auditory). Putting all the analytics into a visual-only format also means more time is needed to process the various graphical information [39], which can lead to fatigue. New modes of insight representations (in addition to visual-only) will elevate the input-channel competitions and ease cognitive processing, in order to more effectively communicate the ever increasing amounts of insights.

More Efficiently Present Insights from Analytics

Given that sound is one of the main channels of communication for humankind, it would be worthwhile for researchers to further explore the use of auditory stimuli in data and insight representations. What we have described in this study is just the beginning. More research will be necessary to understand how information audialization may be used to better represent information for understanding.

There may be merits in revisiting older research on audification and sonification of data to see if some of the methods and approaches are useful in improving the representation of insights through audialization for communications.

More Effectively present Insights from Analytics

Although the process of audialization in Serious Games analytics seems simple in practice, its implications can be quite far-reaching. Some questions that immediately come to mind include: "What range of musical notes would best depict the expertise index?" "Should the similarity indices be depicted using just 10 musical notes – i.e., each note representing 10% of increment in competency? And if so, which 10 notes should be chosen?"

Future research can examine how best to audialize insights from analytics because there are just too many parameters that are subjectable to modification [28]. These include the frequency, duration, amplitude, phase of oscillation, left-to-right channel separation, surrounds, instrument sounds, synthesized vs. sampled audios, etc.

Another area for future research could look into the application of audialization in the emerging field of synesthesia [40] or ideasthesia [41] to enable the perception of (spatial) insights through sound or music.

Can We See What We Hear?

We can all appreciate how the entertainment filming and gaming industry modify our emotions through the use of sounds or music, and enhance our affective experience in those virtual environments. The gaming industry, in particular, has employed various tweets and jingles to represent success and failure in beating bosses, to add to the excitement during battle, to announce the arrival of a hero/villain, and others. Research is currently underway to adapt virtual environments for medical pain management in burnt victims and new amputees to allow them to 'see' and feel something that is not physically real through synesthesia. Sensory substitution researchers [42], [43] are experimenting with sound to convert physical environments into 'soundscapes' to assist the visually impaired to 'see' their environments.

Could a deeper understanding of the audialization process allow human to start to appreciate, if not perceive, the information or insights by forming numeric distances and dimensional spaces in their mind? Could the audialization research contribute to the sensory substitution in human cognition and unlock the neurological pathways to enable people to 'see what they hear' [42]? It is conceivable that these findings could become useful to the field of game design and audialization of insights.

As designers learn to create 'soundscapes' and blend them into the gaming environments, it may give life to new serious games for the visually impaired, for affective gaming, for virtual/sensory reality, and so on. Besides the audialization of existing simulations and serious games for rehabilitation or training, new serious games with audialization capability can present new way to capture analytics and even more insights for visualization, thereby bridging audialization back into serious games.

REFERENCES

1. D.L. Kirkpatrick and J.D. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, 3rd ed. San Francisco, CA: Berrett-Koehler Publishers, 2006, p. 379.
2. H. Ben Zur and S.J. Breznitz, "The effect of time pressure on risky choice behavior," *Acta Psychol. (Amst)*, vol. 47, no. 2, pp. 89–104, Feb. 1981.
3. R. Pieters and L. Warlop, "Visual attention during brand choice: The impact of time pressure and task motivation," *Int. J. Res. Mark.*, vol. 16, no. 1, pp. 1–16, Feb. 1999.
4. M.E. Young, S.C. Sutherland, and J.J. Cole, "Individual differences in causal judgment under time pressure: Sex and prior video game experience as predictors," *Int. J. Comp. Psychol.*, vol. 24, no. 1, pp. 76–98, 2011.
5. C.D. Wickens, A. Stokes, B. Barnett, and F. Hyman, "The effects of stress on pilot judgment in a MIDIS Simulator," in *Time Pressure and Stress in Human Judgment and Decision Making*, O. Svenson and A.J. Maule, (Eds.), Boston, MA: Springer US, 1993, pp. 271–292.
6. K.A. Ericsson, N. Charness, P.J. Feltovich, and R.R. Hoffman, (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. New York, NY: Cambridge University Press, 2006, p. 901.
7. N. Karelaia and R.M. Hogarth, "Skill, luck, overconfidence, and risk taking," Rochester, NY, 2008.
8. S.E. Dreyfus and H.L. Dreyfus, "A five-stage model of the mental activities involved in directed skill acquisition," Berkeley, CA, 1980.
9. S.E. Dreyfus, "The five-stage model of adult skill acquisition," *Bull. Sci. Technol. Soc.*, vol. 24, no. 3, pp. 177–181, Jun. 2004.
10. W.E. Winkler, "The state of record linkage and current research problems," Washington, DC, 1999.
11. M. Bilenko and R.J. Mooney, "Adaptive duplicate detection using learnable string similarity measures," in *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2003)*, 2003, pp. 39–48.
12. C.S. Loh and Y. Sheng, "Performance metrics for serious games: Will the (real) expert please step forward?" in *Proceedings of CGAMES'2013 USA*, 2013, pp. 202–206.
13. C.S. Loh and Y. Sheng, "Measuring the (dis-) similarity between expert and novice behaviors as serious games analytics," *Educ. Inf. Technol.*, vol. 20, no. 1, pp. 5–19, 2015.
14. C.S. Loh and Y. Sheng, "Maximum Similarity Index (MSI): A metric to differentiate the performance of novices vs. multiple-experts in serious games," *Comput. Hum. Behav.*, vol. 39, pp. 322–330, 2014.
15. R.E. Clark, "Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses," *Educ. Technol. Res. Dev.*, vol. 33, no. 4, pp. 249–262, 1985.
16. R.E. Clark, "Media will never influence learning," *Educ. Technol. Res. Dev.*, vol. 42, no. 2, pp. 21–29, 1994.

17. N.B. Hastings and M.W. Tracey, “Does media affect learning: where are we now?” *Tech Trends*, vol. 49, no. 2, pp. 28–30, 2004.
18. C.S. Loh, A. Anantachai, J.H. Byun, and J. Lenox, “Assessing what players learned in serious games: In situ data collection, Information Trails, and quantitative analysis,” in *Proceedings of CGAMES 2007 USA*, 2007, pp. 10–19.
19. G. Zoeller, “Game development telemetry in production,” in *Game Analytics: Maximizing the Value of Player Data*, M. Seif El-Nasr, A. Drachen, and A. Canossa, (Eds.), London, UK: Springer London, 2013, pp. 111–136.
20. A. Drachen, C. Thurau, R. Sifa, and C. Bauckhage, “A comparison of methods for player clustering via behavioral telemetry,” in *Proceedings of the 8th International Conference on the Foundations of Digital Games (FDG 2013)*, 2013, pp. 245–252.
21. C.S. Loh, “Information Trails: In-process assessment of game-based learning,” in *Assessment in Game-based Learning: Foundations, Innovations, and Perspectives*, D. Ifenthaler, D. Eseryel, and X. Ge, (Eds.), New York, NY: Springer, 2012, pp. 123–144.
22. C.S. Loh, “Designing Online Games Assessment as ‘Information Trails,’” in *Games and Simulation in Online Learning: Research and Development Frameworks*, D. Gibson, C. Aldrich, and M. Prensky, (Eds.), Hershey, PA: Idea Group, Inc., 2006, pp. 323–348.
23. C.S. Loh, “Improving the impact and return of investment of game-based learning,” *Int. J. Virtual Pers. Learn. Environ.*, vol. 4, no. 1, pp. 1–15, 2012.
24. J. Velasquez, H. Yasuda, and T. Aoki, “Combining the web content and usage mining to understand the visitor behavior in a web site,” in *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM 2003)*, 2003, pp. 669–672.
25. C. Bauckhage, R. Sifa, A. Drachen, C. Thurau, and F. Hadiji, “Beyond heatmaps: Spatio-temporal clustering using behavior-based partitioning of game levels,” in *Proceedings of the 2014 IEEE Conference on Computational Intelligence and Games*, 2014, pp. 1–8.
26. G. Kramer, “Auditory Display: Sonification, Audification, and Auditory Interfaces,” Sep. 1993.
27. F. Dombois, “Using audification in planet seismology,” in *Proceedings of the 2001 International Conference on Auditory Display (ICAD ’01)*, 2001, pp. 227–230.
28. T. Hildebrandt, S. Kriglstein, and S. Rinderle-Ma, “On applying sonifications methods to convey business process data,” in *Proceedings of the CAiSE’12 Forum at the 24th International Conference on Advanced Information Systems Engineering (CAiSE Forum)*, 2012, pp. 74–81.
29. P. Yalla and B.N. Walker, “Advanced auditory menus,” 2007. Online. Available: http://sonify.psych.gatech.edu/publications/pdfs/2007-GVU0712-YallaWalker.pdf Accessed: 12-Jan-2015.
30. A. Cox, “Fractions of a second: An Olympic musical,” *The New York Times*, 2010.
31. P. Yalla and B.N. Walker, “Advanced auditory menus: Design and evaluation of auditory scroll bars,” in *Proceedings of the Tenth International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2008)*, 2008, pp. 105–112.
32. C.S. Loh, Y. Sheng, and I-H., Li, “Predicting expert-novice performance as Serious Games Analytics with objective-oriented and navigational action sequences,” *Comput. Hum. Behav.*(in press).
33. G. Vettigli, “Sound synthesis with Numpy,” 2011. Online. Available: http://python.dzone.com/articles/sound-synthesis-numpy. Accessed: 12-Jan-2015.
34. L.E. Nacke and C.A. Lindley, “Flow and immersion in First-Person Shooters: Measuring the player’s gameplay experience,” in *Proceedings of the 2008 Conference on Future Play*, 2008, pp. 81–88.
35. L.E. Nacke, C. Lindley, and S. Stellmach, “Log who’s playing: Psychophysiological game analysis made easy through event logging,” in *Fun and Games*, vol. 5294/2008, P. M. E. AL, Ed. Springer Berlin/Heidelberg, 2008, pp. 150–157.
36. C.S. Loh, Y. Sheng, and D. Ifenthaler, (Eds.), *Serious Games Analytics: Methodologies for Human Performance Measurement, Assessment, and Improvement*. New York, NY: Springer, 2015.
37. A. Canossa, M. Seif El-Nasr, and A. Drachen, “Benefits of game analytics: Stakeholders, contexts and domains,” in *Game Analytics: Maximizing the Value of Player Data*, M. Seif El-Nasr, A. Drachen, and A. Canossa, (Eds.), London, UK: Springer London, 2013, pp. 41–52.
38. C.D. Wickens, “Multiple resources and performance prediction,” *Theor. Issues Ergon. Sci.*, vol. 3, no. 2, pp. 159–177, 2002.
39. A. L. Dyckhoff, D. Zielke, M. Bültmann, M. A. Chatti, and U. Schroeder, “Design and implementation of a learning analytics toolkit for teachers,” Educ. Technol. Soc., vol. 15, no. 3, pp. 58–76, 2012.
40. K.S. Suslick, “Synesthesia in science and technology: More than making the unseen visible,” *Curr. Opin. Chem. Biol.*, vol. 16, no. 5–6, pp. 557–63, Dec. 2012.
41. E. Milan, O. Iborra, M.J. de Cordoba, V. Juarez-Ramos, M.A.R. Artacho, and J.L. Rubio, “The Kiki-Bouba effect: A case of personification and ideaesthesia,” *J. Conscious. Stud.*, vol. 20, no. 1–2, pp. 84–102, 2013.

42. A. Haigh, D.J. Brown, P. Meijer, and M.J. Proulx, "How well do you see what you hear? The acuity of visual-to-auditory sensory substitution," *Front. Psychol.*, vol. 4, p. 330, Jan. 2013.

43. J. Carmichael, "Device Trains Blind People to 'See' by Listening" *Popular Science,* 2013. Online. Available: http://www.popsci.com/science/article/2013-07/synesthesia-blind. Accessed: 10-Feb-2015.

Citation: Loh, C. S. & Ekstrand, A. (2015). *Audialization in Serious Games Analytics: Visualizing player performance improvement by sound or music.* Proceedings of the 3rd International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2015), satellite of International Conference on Intelligent User Interfaces. (March 29, 2015) Atlanta, GA.