Chapter 3
A Conceptual Framework for Gamification Measurement

Ronald Dyer

3.1 Introduction

This chapter's focus is on contributing to the body of knowledge as it relates to the measurement of gamification outcomes, i.e. a methodology for the assessment of gamification implementation within organizations. The term gamification has come to the forefront with much fanfare and receptivity by both educators and corporate training professionals. One only has to look online to see the myriad number of conferences, publications and blogs devoted to the topic and perpetuation of its perceived benefits. Figure 3.1, below indicated the number of titles on the topic illustrating continuous increases in the subject matter over the last four years. Yet despite the large amount of hits on the topic, there still remains a lack of coherent understanding on what kinds of studies and results gamification has yielded (Hanari, Koivisto, & Sarsa, 2014). Moreover, understanding the effectiveness of gamification remains a pertinent issue.

The term gamification as defined by Huotari and Hamari (2012a, 2012b) is as follows:

> A process of enhancing a service with (motivational) affordance in order to invoke gameful experiences and further behavioral outcomes.

As such the use of gamification involves the utilization of several of the concepts inherent in games such as level design, tokens, badges and other rewards to incentivize learning during play. What makes the use of gamification unique when pitted

R. Dyer (✉)
Grenoble Ecole de Management, 12-Rue Pierre Semard,
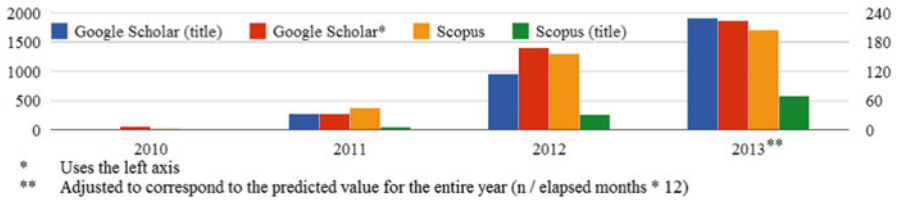Grenoble France 38000
e-mail: ronald.dyer@grenoble-em.com

*   Uses the left axis
**   Adjusted to correspond to the predicted value for the entire year (n / elapsed months * 12)

Fig. 3.1 Search hits for "gamification"

against traditional learning and development methodologies both in education and business is the element of play. According to Weisberg, Hirsh-Pasek, and Golinkoff (2013) play is ubiquitous and linked to positive social behaviours. Hirsh-Pasek, Golinkoff, Berk, and Singer (2009) building on the works of Vygotsky (1978) and Piaget (1962) further state that "play is not only important for encouraging the development of socially relevant skills like self-regulation and empathy but also assisting in both academic and social learning," Weisberg et al. (2013). As with most learning frameworks play has long been considered *non-relevant* as an engagement factor for fostering improved cognition within education. However, recent advanced in the areas of game-based learning and serious games respectively have demonstrated clearly that this is no longer an applicable approach and that play has an integral role in learner retention and engagement.

While we herald the up-surge of gamification as a *tour de force* in education and business there remains a fundamental issue which needs to be addressed, i.e. measurement. For all the hype associated with gamification, for buy-in and adoption continuity to persist there must be robust assessment and measurement of its outcomes if communities of practice are to successfully sustain present momentum. Failure to provide empirical approaches to measurement of gamification will eventually result in dissipation of its relevance and at best relegation to the archives of education and business as another fad.

The question becomes, how do we measure the benefits of gamification? The use of traditional metrics in their existing format cannot be as easily applied to this field as historical assessment of student performance has focused on teacher evaluations, with testing as one of the primary instruments. Gamification caters to a cadre of learners dubbed "Millennials" or the Net Generation, Oblinger (2004) whose preferences tend towards teamwork, experiential activities, structure and the use of technology. Traditional pedagogical practices are still constrained by their traditionalist nature/thinking with the teacher as leader and the student as passive recipient, as such they require significant re-tooling to even attempt preliminary assessment(s). On the other hand, few if any measurements are readily available to would be educators and trainers wishing to assess learner performance specifically utilizing gamification.

The chapter provides insight on the integration of a proposed measurement framework for gamification into organizations (new and existing) to support successful implementation and sustainable adoption of gamification based approaches.

The proposed outcome of this chapter from the author's perspective will be a measurement framework for educators and business persons alike charged with responsibility for integration to better approach the area and provide clear sign-post to guide them in their individual efforts as practitioners contributing to the learning and development ecosystem.

3.2 What is Gamification

Deterding, Dixon, Khaled, and Nacke (2011), provide a further definition of gamification as "the use of game design elements in non-game contexts." In tracing the historical context of gamification and its place in the digital media industry its emergence dates back to 2008. The popularity effectively commenced around 2010 and has since 2010 managed to institutionalize itself as a common household term. How did we arrive at a gamified universe? From a conceptual perspective the notion that user interface design can be informed by other design practices has its roots in the Human Computer Interaction (HCI) movement. Carroll (1982) analysed the design of early text adventures leading him to suggest research programs on fun and games in relation to ease of use.

Deterding et al. (2011) indicated that the maturation and expansion of the field has led to further research into areas like motivational affordance and pleasurable products. These approaches take into consideration the aspirational phenomenon that is today branded gamification. Gamification falls into the potential category of games with a purpose in which game-play is piggybacked upon as a means to solve human tasks. It is within this context that the "playfulness" aspect of games has given rise to a gamification movement as a mode of interaction inculcating the aspects of game mechanics. In its present form gamification demarcates itself through complexity, given the delicate balancing act of game design and motivational theories which provide the aspects of gamefulness, gameful interaction and gameful design all requisite elements of play. These aspects feed into the definition provided above and differentiate gamification from play. Gamification relates to games not play, where play can be conceived as the broader, looser category containing different elements of games, Caillois (1961). Gamification is predicated by structured rules and competitive strife towards goals. Hence the term gamification and its linkages to the concepts of:

1. **Gamefulness (the experiential and behavioral quality)**
2. **Gameful Interaction (artifacts affording that quality), and**
3. **Gameful Design (designing for gamefulness, typically by using game design elements).**

These terms/concepts in defining gamification provide a clear distinction between games and play/playfulness and as such attributes to the novelty that is gamification. Gamification can inherently give rise to playfulness but the inverse does not necessarily

apply given the requisite structures and mindset required for the former? Thus gamification applies to the incorporation of elements of games, Braithwaite and Schreiber (2008) and as such places consideration on the artifactual and social elements of games without delving deeper into the software application element more akin to game studies etc. In essence a game is rule based, where outcomes are assigned varying values and the player exerts effort in order to influence the outcome. From a terminology perspective, the term gamification gains better clarity when associated with game-design and not the wide game eco-system associated with game-based technologies. Therefore, in situating gamification and its role, we must understand its use in relation to an extension of design, (not within the context of technology), the elements requisite for play (vs. a full game), characteristics of the game rather than playfulness and its non-game context regardless of usage, context or media associated with implementation. This broader context allows for an understanding of the workings of gamification within the broader framework of ludification. As such its role as a lever in playfulness' integration within the larger game eco-system such as game-based learning and serious games is critical.

There is however another school of thought on the gamification definition as presented by Huotari and Hamari (2012a, 2012b). They define gamification as "service packaging where a core service is enhanced by rule-based service systems that provides feedback and interaction mechanisms to the user with an aim to facilitate and support the users overall value creation." This definitional approach highlights the goals of gamification from an experience perspective and as such move away from the notion that it is purely based on game elements. However, gamification is not always executed within the context of concrete game elements but is more of a process. As such gamification is supposedly imbued with gameful experiences which seek to improve service delivery within an affordance context. Huotari and Hamari (2012a, 2012b), referring to affordance here as *any qualities of the service which contribute to the emergence of the gameful experience*. This is where we can now introduce the nexus of our discussion on the benefits of measurement of gamification as current thinking on the subject seeks to eke out a measurement context based on a formulary of sales increases, clicks and general learner retention. The benefits of gamification cannot be measured on the axis of these metrics nor can we expect the integration of existing theories of learning to provide a cookie cutter measurement ethos. Let us commence our measurement discussion by first examining three of the most widely utilized learner evaluation metrics and their value within a gamification context.

3.3   Traditional Measures of Learning

When we think of learning within the parameters of assessment the familiar approach of assessment rubrics, portfolios and other examination-based methods immediately come to mind. What all of these approaches have in common is the use of the same underlying theories or schools of thought as it relates to how to measure

the progress of learners and or/individual improvement. In the field of learning there are three theories that standout:

1. *Kirkpatrick's Four Levels of Evaluation Theory*
2. *Kolbs' Experiential Learning Theory and*
3. *Sweller's Theory on Cognitive Load*

Each of these has earned their rightful place in the arena of learning and are amongst the most widely used evaluative metrics in the field. While their value cannot be underscored the question remains as to their applicability to an emerging field such as gamification. Based on present trajectory, gamification represents a hybrid approach to learning reinforcement and retention and given the long history of these three theoretical approaches may be challenged in measurement as a result. If gamification is to be taken much more seriously, especially to convince its detractors there needs to be either the development of new measurement theories or modification of existing to allow for better evaluation of the benefits in line with further adoption. Let us first examine these three underlying theories and the potential value that may exist for integration into gamification measurement.

3.4 Kirkpatrick's Four Levels

The Kirkpatrick four-level framework is one of the most widely utilized evaluative models for training employees specifically with sales force practitioners. Originally developed by Donald Kirkpatrick (1994) cited by Tan and Newman (2013) it consist of four evaluation levels (viz., reaction, learning, behavior and results) arranged in ascending order and varying degree of difficulty to accomplish. Each level seeks to measure or evaluate as follows:

Level 1—Reaction: Measures how participants have reacted to the training
Level 2—Learning: Measures what participants have learned from the training
Level 3—Behaviour: Measures whether what was learned is being applied on the job
Level 4—Results: Measures whether the application of training is achieving results

The levels are layered in such a way that the success of one builds on the previous adding to the precision of measurement. Delving further into each level, their value lies firstly in the learner's perception and reaction towards a training event and the benefit of learning (positive or negative) that occurs. Secondly, it measures what advanced knowledge has been gained from training and the extent to which skills/attitudes are impacted. Thirdly, it focuses on what was learnt and whether is it being applied on the job. The third level focuses strongly on knowledge transfer specifically in-line with real world circumstances. Finally, level four examines the application of training in achieving quantifiable financial results. It careful examines the success of training in relation to increases in production, sales, decreased costs and improved quality etc. requisite for positive return on investment (ROI).

This framework augurs well for the potential scope of gamification as it has clear linkages with trainee performance and effectiveness to accentuate the importance of aligning training interventions with gamefulness and by extension the strategic focus of performance improvement in organizations. The evaluation of gamification needs to start with a rigorous process of data collection in order to analyze its effectiveness and benefits within a Return on Investment (ROI) context. Kirkpatrick's approach possesses some of the requisite elements to afford such an evaluation via its levels. The question remains as to whether the impact on gamification would be formative (focused on the actual process of training/performance improvement) or summative (focusing on the final product or result of the process), Rahimic and Vuk (2012) and the value proposition which this approach would provide to assessors. Gamification represents a form of tool and as such, justification of the cost and effectiveness for sustainability requires a thorough analysis to ascertain the impact on training design and ultimately the possible benefits for enterprise. At the cursory level, Kirkpatrick's framework brings to gamification some components of a scorecard model with the requisite elements for evaluation of gamification generating evaluative criteria along six potential elements:

1. *Reaction*
2. *Learning*
3. *Application/Impact*
4. *Organizational Impact*
5. *Return on Investment and*
6. *Non-quantifiable benefits*

Leveraging these foundational elements of Kirkpatrick's model into gamification measurement can provide some clues as to the linkage between the design/game mechanics components and training results by collecting data within the framework of the (6) defined elements as proposed. However, this is but only one of the theories on learning evaluation and we must examine the others prior to formulation of a proposed framework for measuring gamification in a holistic way. We shall now take a look at Kolbs' Experimental Theory.

3.5   Kolbs' Experiential Theory

Experiential Learning Theory (ELT), Kolb (1984) has its roots in the work of John Dewey (philosopher and psychologist), Kurt Lewin (psychologist) and Jean Piaget (psychologist). It is intended to provide a holistic adaptive process on learning which merges experience, perception, cognition and behavior, citing McCarthy (2010). ELT defines learning as "the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience" Kolb (1984). The learning model is cyclical in process consisting of four components, concrete experience (CE), reflective observation (RE), abstract conceptualization (AC) and active experimentation (AE).

The learner must continuously choose which set(s) of learning abilities to use within specific learning situations and successfully perceive new information through the experiences. These experiences can be concrete, tangible and felt or abstract conceptualizations bearing in mind that learners can enter the model at any stage. Given that the learning style preference akin to this model tends to lend towards perception and grasping of new information through symbolic representation (thinking about, analyzing or systematically planning), how/where does gamification fit in from a measurement perspective? The answer to that questions lies somewhere in the Learning Styles Inventory (LSI) citing McCarthy (2010) an instrument used to assess the individual learning styles, identifying four types of learners based on their approaches to obtaining knowledge through one or more of the following methods:

Divergers: Those who prefer to approach learning through concrete experience (CE)
Accommodators: Those who prefer to learn hands-on
Assimilators: Those who prefer to approach knowledge through abstract conceptualization and
Convergers: Those who prefer to approach knowledge through active experimentation

The development of Learning Styles Inventory (LSI) Kolb (1971) as cited by McCarthy (2010) provides an individual understanding of the learning process through experience and their individual approach to learning. Therein lays the potential value to gamification as it represents a unique baseline for not only how individuals learn best but serves as an investigative tool into specific characteristics of individual learning style infused with gamification methodologies. Hence, a Kolb based approach to measurement of gamification may provide some preliminary insight between experience and conceptualization as a learning reward system which provides the value added of incentivized reflection post training.

3.6   Sweller's Theory of Cognitive Load

Cognitive load theory (CLT) Sweller (1988, 1989) as cited by Chandler and Sweller (1991) is concerned with the manner in which cognitive resources are focused and used during learning and problem solving. Many learning and problem solving procedures encouraged by instructional formats results in activities by participants far removed from the task at hand or lost soon thereafter post training. This theory's relevance to the topic of gamification and its adoption are important in that cognitive load represents an integral measurement tool for investigating individual performance. Measuring the impact of cognitive load on gamification goes a long way in presenting robust evidence to skeptics whose present thinking on the topic requires further convincing.

The basic premise of (CLT) is reduction of the load for difficult task so that it becomes easier to remember within a working memory context. Cognitive load can

be broken down into three types Intrinsic, Germane & Extraneous Sweller (1999, 2005) and Mayer (2005a); Mayer and Moreno (2003). DeLeeuw and Mayer (2008) explain each as follows: *extraneous processing*, in which the learner engages in *cognitive* processing that does not support the learning objective (and that is increased by poor layout such as having printed words on a page and their corresponding graphics on another page); (b) *intrinsic (or essential) processing*, in which the learner engages in cognitive processing that is essential for comprehending the material (and that depends on the complexity of material, namely the number of interacting elements that must be kept in mind at any one time); and (c) *germane (or generative) processing*, in which the learner engages in deep cognitive processing such as mentally organizing the material and relating it to prior knowledge (and that depends on the learner's motivation and prior knowledge, as well as prompts and support in the lesson). Cognitive load is a valuable measurement variable for facilitating the intrinsic learning reinforcement goals of gamification as they present a strong foundation for learner retention. The effectiveness of training rests in its ability for guided instruction and facilitation of learner's mental integration of disparate sources of mutually referring information. For example such as when instructors use separate text and diagrams to validate or emphasize and participants have to reflect on each element separately. These instruction(s) generate a heavy degree of cognitive load, because material must be mentally integrated before learning can commence. Processing and analysis therefore takes a greater time creating a lag of taught knowledge transfer back into the work environment. This choice of theory rest with prior success of testing experiments Chandler and Sweller (1991) where content of a split-source nature (content & diagrams) and integrated information were designed for comparison using conventional instruction integrated over a period of several months in an industrial training setting. The value here of gamification, is adding through rewards and incentive a further motivation to "get it." Therefore, learning takes on a more focused/commoditized value to the learner through incentives.

Moreover, it can assist in discerning whether the learners are more focused on the reward/incentive than gestation of knowledge. Choi and Lee (2009), applied the concept of cognitive load to enhance student's problem solving abilities as well as assist in implementing a model to improve teacher education with students' real-world problem solving abilities. The result was the use of this theoretical approach to provide a framework to deal with dilemmas faced by practicing teachers in elementary classrooms. If an incentive-driven component via gamification is added then perhaps the approach can provide richer results more in-line with results anticipated in real-life situations as life decisions are very much incentivized.

Gamification holds potential to support O'Neil, Baker, and Wainessa (2005) theory on the effectiveness of game environments documentation in terms of intensity and longevity of engagement as well as their commercial success. It allows for the capturing of performance variables that would not normally be tracked but do add value to the overall learning process. However, since there is much less solid information about which outcomes are systematically achieved using individual and multiplayer games to train participants in acquiring knowledge and skills gamification's performance imperative remains questionable. The lack of clearly presented

measurement metrics required to measure the degree to which their design fosters the desired knowledge and skills transference is still debatable. Secondly, the impact of game play needs to be studied to determine what works. Cognitive load measurement within a gamification context allows for evaluation and accumulation of clear evidence of impact. Failing to apply robust measurement metrics as stated can result in a tendency to dismiss game environments and gamification as motivational fluff.

3.7   Towards a Unified Framework for Gamification Measurement

The question becomes, given the challenges of a robust system to effect measurement, how do we find a way to integrate existing theories of learner evaluation to provide a comprehensive system of metrics? It is unlikely at this time given the relative novelty of gamification that a complete system of measurement will arise in the near future. However, there exist more than enough measurement tools between gaming and learner development to create a hybrid methodology, which may provide better metrics for overall assessment.

To develop such a framework we need four key ingredients that will assist in assessing each stage of the gamification process and provide an integrative model. They are as follows:

1. *Play Assessment Diagnostic*
2. *A Gamification Scorecard*
3. *Pre/Post Knowledge Assessment*
4. *Gamification Performance Assessment Review*

3.8   Play Assessment Diagnostics

In order to better understand the context within which gamification can best be measured readers must first understand how to assess play specifically within androgogical environments. Social development is an integral component of early childhood and similary are critical to the ability to interact inter-organizationally in later life. Play assessment represents a way to identify which students/employees may be developmentally behind their peers and as such at risk. In essense it makes screening students/employees easier, time efficient, and cheaper through comprehensive interaction in a "live" environment. Like many other game environments and formats of play, gamification assessment is rooted in observation. While much data can be gathered via the process of game analytics systems they may not tell the entire story. As such measurement of gamification requires a caliber of metrics that are both data driven and observational. As a first stage in the gamification measurement process, we must first observe and measure the element of play interaction. One of the better ways to approach this is through the utilization of stealth assessment.

Stealth assessment is woven directly into the fabric of instructional environments to support learning of important content and key competencies, Shute (2011). The power of stealth assessment lies in the performance data gathered through the course of play/learning and the inferences which can be made Shute, Ventura, Bauer, and Zapata-Rivera (2009)). To measure gamification effectively we must create an environment of inference which are stored in the dynamic models of learners and are a direct function of *flow* i.e. the state of optimal experience, where a person is so engaged in the activity at hand that self-consciousness disappears, sense of time is lost and the person engages in complex, goal-directed activity not for external reward, but simply the exhilaration of doing, Csikszentmihalyi (1990). The key elements of the approach include:

- **Evidence-Centered Assessment: which systematically analyses the assessment argument concerning claims about the learner and evidence that supports those claims, Mislevy, Steinberg, and Almond (2003) and**
- **Formative Assessment: a range of formal and informal assessment procedures employed during the learning process in order to modify teaching and learning activities to improve student attainment.**

The power of stealth within gamification rest in rethinking assessment not linked to the world of multiple choice answers but to the identification of new skills and standards and measuring them relevant to the twenty-first century. As such given that gamification is not a standardized academic or educational learning process embedding stealth assessment has great potential to increase learning given its existing history in game-based environments. This form of measurement for gamification presupposes two underlying assumptions, (1). learning by doing improves the process and its outcomes and (2). different types of learning and learners attributes may be verified and measured during game-play that would not otherwise be captured. So how would stealth assessment work to inform play assessment diagnostics?

Given that there remain a challenge to educators/trainers who desire to employ gamification and other game-design elements to support learning an evidence-centred design (ECD) process: Mislevy et al. (2003) is required for effectively measuring student retention and competence. The fundamental idea behind ECD originates from Messick (1994)) and formalized by Mislevy and Haertel (2006). The process commences with an identification of what knowledge, skills or other attributes should be measured. These variables are observed directly so behavior and performance can be aligned to the overall capture of relevant data. This should be followed by a determination of tasks/situations which would draw out such behaviours/performance. The nature of gamification lends to this process aptly as gamification elicits behaviours which bear evidence about key skills and knowledge through demonstration of the following as identified by Shute et al. (2010):

1. **What collection of knowledge and skills should be incentivized within an assessment context? That is Competency Models (CM), which supports grading, certification and diagnostic for further instructional support.**

2. **What behaviours or performance should reveal the relevant constructs associated with a reward system? That is Evidence Models (EM) which expresses how learner's interaction and responsiveness to a given problem constitutes evidence about their competence and by extension receipt of a reward for positive performance.**
3. **What task should elicit the behaviours which comprise the evidence, i.e. (TM)? This component provides a framework for the construction of situations which provide interaction evidence targeting aspects of knowledge related to the specific competence(s).**

However, there remains an additional component of the stealth process' induction into gamification measurement that provides the robust value added to ensure measurement effectiveness. Bayesian networks, Pearl (1988) cited by Shute (2011). These networks provide a useful model to handle uncertainty by using probabilistic inference to update and improve the data regarding learner competencies and gamification effectiveness. By using what-if scenarios in the observance of evidence which describes particular situations from a predictive perspective coupled with the ECD view allows propagation of information on the achieved behaviours. The Bayesian approach allows a comparative between the what-if and what-is so that the resulting probabilities inform future decision making on which elements of gamification work, where "*incentivization*" should be focused and selection of the best content chunks which had relevant impact from a delivery perspective. The combination of this built-in play assessment diagnostic into the gamification measurement framework permits all measurement stakeholders to examine the evidence/success of gamification under the lens of approximate competency levels (present & future state). It also has potential for validation of what the student/participant can do without disruption of the play experience and consequently their flow.

3.9   A Gamification Scorecard

The complexity of gamification measurement requires more than one system in place to effectively analyse the impact of gamification interventions. As such to effectively collect data from the play assessment diagnostics we require a requisite instrument. The Game Performance Assessment Instrument in the author's opinion represents a unique opportunity for modification, capture and representation of data on gamification performance and effectiveness.

The Game Performance Assessment Instrument (GPAI), Oslin et al. (1998) cited by Memmert and Harvey (2008) was developed to measure game performance behaviours that demonstrate tactical understanding, as well as player ability to solve tactical problems by selecting and applying the appropriate skills. It is most commonly used with assessment of physical sports but the author felt that given its robust performance metrics and ability to identify nonspecific observation components it can potentially be expanded for gamification measurement. GPAI examines such

Table 3.1 Game components observed in GPAI—*Source*: (Memmert & Harvey, 2008)

Game component	Description
Decision Making	Player makes appropriate decisions about what to do during the game
Skill Execution	Player efficiently executes selected skills
Adjust	Player movements offensive or defensive are necessitated by the flow of the game
Cover	Player provides appropriate defensive cover to help, backup or challenge opponents
Support	Player provides appropriate support to teammates
Guard/Mark	Player engages appropriate strategy(s) to ward off opponents who may threaten present position
Base	Player appropriately return to a recovery position between skill attempts

tactical components as base, adjust, decision(s) made, skill execution, support, guard/mark and cover (i.e. assistance & support to team) all of which have a tremendous impact on game play performance and requisite elements of evaluation in gamification. Table 3.1, below provides a detailed description of each tactical element.

Depending on the degree of gamification involved assessors can select one or more categories for evaluation of performance. The benefits of this approach to a gamified environment being (a) it can be adapted to various activities and (b) it has the ability to not only measure psycho-motor skills but skills which would not be effectively measured using traditional approaches, Mitchell et al. (2006) cited by Memmert and Harvey (2008). These would include skills such as adjustment and cover as it relates to team performance all of which are requisite components for gamification activities. From a measurement perspective the process also lends well to gamification through its scoring method utilizing a tally method based on a Likert scale, Memmert and Harvey (2008). The use of such a scale would provide categorical descriptors which provide easier metrics for performance scoring as well as a wider variety but not so wide that reliability of measurement would become difficult. An example of a rating system based on GPAI for gamification could look something like the example listed below:

The value of implementing such a system in gamification allows for both peer-review of performance as well as codification of game-mechanic components using an effective tally system. Furthermore, components of GPAI (such as decisions made, skill execution, support and adjust) approaches have already been validated as effective in some physical sports and as such form the basis of a framework with a track record which can be modified for gamification assessment. The use of this coding system allows each coder responsible the ability to individually observe behaviours, assess as appropriate/effective or inappropriate/ineffective then tally responses and create indices for decision making across a single component or multiple environments Memmert and Harvey (2008). The impact of tallying these components on an index for gamification measurement can then be disaggregated into two index measures game performance (GP) and game involvement (GI) and effectively

correlated to provide an overall assessment of the gamification experience. While there are inherent limitations of GPAI such as calculation of individual and overall game performance indices, use of game involvement vs. game performance index, observer reliability, nonlinearity and finally usefulness of action there is value to the approach. A closer examination of learning time using GPAI both before during and after data coding will result in more stringent levels of observation reliability especially if two independent observers/coders are utilized thus reducing the problem of nonlinearity. Furthermore, it reinforces the GPAI validation process and reduces the level of subject biases and uselessness of action through an assessment planning cycle (the before, during and after process) that makes the gamification process more congruent with the overall learning objectives.

3.10   Pre and Post Knowledge and Skills Assessment

The success of any learning intervention hinges on the ability to assess participant's knowledge and skill both at the beginning and end of a process. The value of this methodology is to design effective intervention to close the relevant learning gaps. This methodology's perspective, for pre and post knowledge/skills assessment either summative or formative assist learning designers/delivers to better ascertain course effectiveness.

  We have explored both play assessment diagnostics and GPAI; we now add a third component for measuring gamification the knowledge assessment. According to Balakrishnan, Bengasamy, and Aman (2011) traditional skills approaches are widely used in teaching games as the direct instructional method citing Metzler (2000). The emphasis has primarily been on skills and drills assessment without a clear consideration for the games themselves. Given that learning is an active discovery activity with learners engaged in construction of tactical understanding an assessment process needs to have clear sign post as it relates to its perceived problem solving and decision making activities. As such for gamification to be successful and the determinants of its success measurable, a pre/post assessment component should be integrated. Therefore the learning environment inculcated through the gamification design process must be identified and assessed prior to implementation. Additionally, prior to implementation of gamification assessment parties should first understand and gauge:

1. **The player(s) depth of understanding of the knowledge to be gained through the gamification experience. That is such factors as the degree of metacognition, and the conditions under which knowledge would be utilized.**
2. **The meaningful concepts the player(s) understands within the content domain. That is strategic knowledge which are memorized and used either to solve problems or effect new knowledge.**
3. **The degree to which the player(s) is already transferring knowledge amongst other members of their existing team(s). That is assessment of transactive**

memory responsible for the encoding, storing, retrieving and communication of group knowledge and develops over time as group members communicate, Lewis, Lange, and Gillis (2005).

Hence, a constructivist approach is required where an active learning approach takes place with participants personally constructing and interpreting the information based on their experiences, which will consists of the eventual gamification output. Moreover, given that constructivism is a participant-centred approach based on learner's perspectives (single & multiple) there is a need to assess these realities both prior to and upon completion of the gamification experience. The pre/post assessment process must consider the development of understanding the learning activities as identified by Fig. 3.1 below:

What the above figure tells us is that as participants confront new, unfamiliar features of their environment which do not fit with their existing view of reality Piaget and Inhelder (1969) cited by Balakrishnan et al. (2011) a disequilibrium occurs. The author suggest pre/post assessment to first understand the disequilibrium (pre-assessment), design the gamification experience within this context and post assess to ascertain if the gamification experience fits with the new experiences. Additionally, if their cognitive structures (also associated to Cognitive Load) have changed how do they accommodate these experiences. The entire pre/post assessment experience falls within the constructivist learning constructs allowing participants to engage in activities which require higher level thinking but with designers/assessment professionals having a clear understanding of how to apply gamification to the knowledge building activities. Finally, since participants will try to assimilate all stimuli associated with the mechanics of gamification into their existing schemas a pre/post assessment methodology provides designers/assessors with an answer as to "what do they require to succeed in this situation and if their existing knowledge schemas can sufficiently address the question. The pre/post approach tells us from a constructivist perspective how participants will learn best and connect that process to the design and ultimate measurement of gamification through the combination of prior and new knowledge Fig. 3.2.
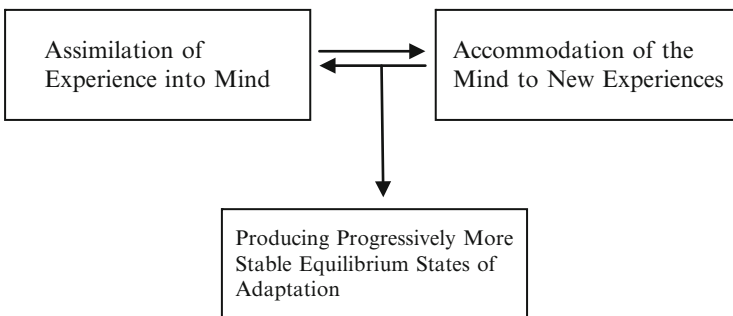

Fig. 3.2 Development of understanding (*Source*: J. Piaget & B. Inhelder, 1969)

3.11   Gamification Performance Assessment Review (GPAR)

The final stage in the assessment process is the Gamification Performance Assessment Review (GPAR). We have explored various measures of assessment which have so far focused primarily on the player(s). There remains however a need to also assess the gamification from a high-level so as to ascertain the effectiveness of gamification as the chosen strategic tool to improve performance of targeted participants. With the growing attention that gamification has received there is also a need to ascertain their value proposition (i.e. the value gained by customers/participants for using the product or service). Thus far gamification is spreading like a wildfire, perhaps to some extent without significant controls. One only has to do a web search and look at the number of hits, new feeds and media events associated with the term. Few theories exist as to an appropriate proposition for a comprehensive framework for evaluation and even fewer operationalized models to examine their causality. The considerable number of publications on the subject of gamification though burgeoning only further propel the hype leaving detractors more suspect and requesting measurable evidence to support its largess. Taking a page from the proposed evaluation framework proposed by Mayer (2012) with some modifications we can ask the following questions of gamification:

1. **What are the requirements and design principles for a comprehensive methodology for its evaluation?**
2. **To what extent does gamification contribute to learning in a real context?**
3. **What are the factors/components in gamification, which contribute to this learning?**
4. **To what extent are the "learnings" purported by gamification transferrable?**

These are all valid questions, which a GPAR analysis akin to a performance audit may possibly answer. Core to the deconstruction of gamification is the need to understand the conditions which make gamification and acceptable intervention, the quality of the intervention, the population demographics, mediating variables and relevant background elements which attribute to the game-mechanics associated with design. A GPAR Analysis therefore requires the following, Fig. 3.3 refers:

1. **Contextualization**: gathering of data as it relates to the special features of gamification, the environment within which it operates and the observations/assumptions which can be made when it is in action. For this contextualization of GPAR to take place there needs to be an underlying hypothesis which entails the design-orientation (artifacts), domain-orientation (the effectiveness of the use of gamification, its complexity and dynamics) and the disciplinary-orientation (the culture, ethics, politics etc.).
2. **Operationalization**: This event needs to occur both pre/post game. At the pregame stage the review will examine such factors as demographics, prior experience with gamification, attitudes towards game-play, pre-existing skills sets (game-play) and group/team characteristics so as to determine its effectiveness. In-game GPAR will examine player-performance, game-play as it relates to
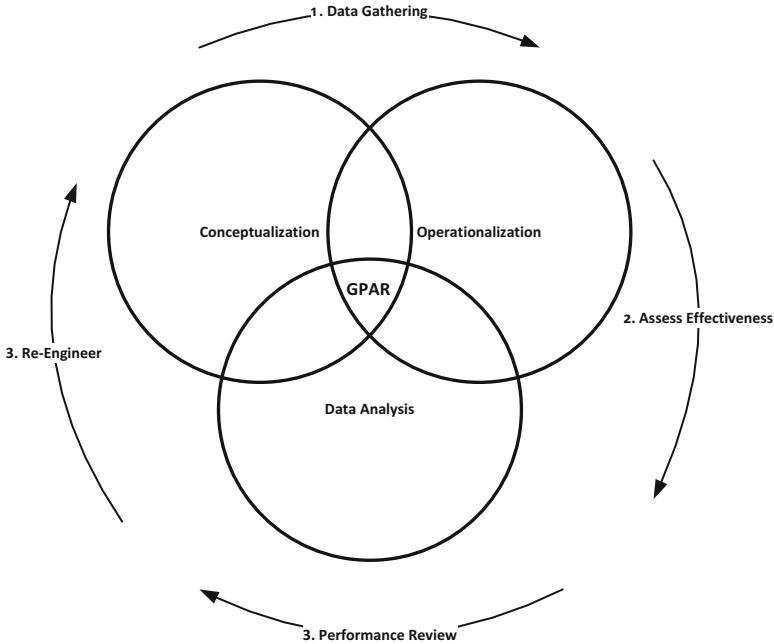
Fig. 3.3 Game performance assessment review process

effort, influence, power etc., and game experience, i.e. flow, immersion and presence. Finally at the post-game stage GPAR will again review the game experience (was it fun/beneficial), degree of player satisfaction, first order learning (individual/participant and short-term in nature), second order learning (longer-term, group, knowledge reconstructions).

3. **Data Analysis**: This is concerned with the analysis of data from the contextualization and operationalized review to ascertain the degree of influence of gamification and distinguish its impact. The analysis will test the overall efficacy of gamification within a group as well as comparatively across several groups.

GPAR is a necessary and final element in the review process as it allows for a strategic view of gamification which does not examine individual player performance but examines the overall performance of the game allowing us to find the influencing factors regarding the efficacy of gamification, improve the gamification constructs for future designs and provide the requisite empirical data to better convince/reinforce adoption Table 3.2.

3.12 Framework Application Process Methodology

To ensure the success of the gamification measurement framework an integrated process approach which brings together all the components must be illustrated i.e.

Table 3.2 Likert scale ratings for gamification assessment

| Rating	Definition
5. Very Effective Performance	Player always attempting to utilize training/learning(s) and communicates learning(s) effectively with teammates
4. Effective performance	Player attempts most times to utilize training/learning(s) and communicates learning(s) effectively most times with teammates
3. Moderately Effective	Player begins to demonstrate communication of training/ learning(s) effectively with teammates
2. Weak Performance	Player rarely utilize training/learning(s) and communicates learning(s) effectively with teammates
1. Very Weak Performance	Player never utilizes training/learning(s) and communicates learning(s) effectively with teammates
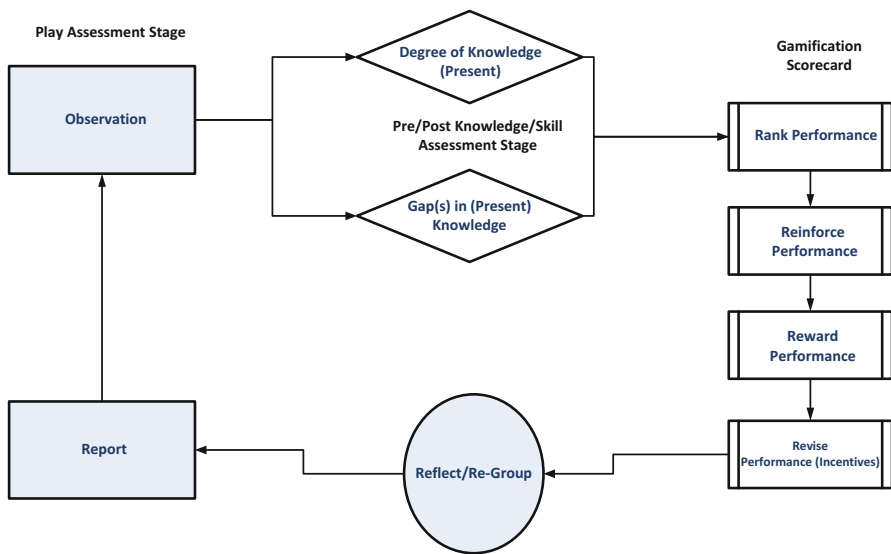
Fig. 3.4 Gamification measurement framework (Vers 1.0)

1. *Play Assessment*
2. *Gamification Scorecard*
3. *Pre-Post Knowledge Assessment*
4. *Gamification Performance Assessment review*

    To better understand the integration of the proposed framework, the following model is proposed to support implementation, Fig. 3.4. At the initial stage, Play Assessment integrators of gamification methodologies must engage in passive observation. The approach allows for ability to gauge and document specific reactions/behaviours inherent with the Gamification methodology and ascertain acceptance. At this stage it is also suggested that integrators keep some form of learning journal which clearly documents player performance. The second stage encompasses assessment of knowledge/skills (pre/post). As indicated earlier in the rele-

vant section this represents a discovery process and as such the aim is to understand the depth of knowledge/skill associated with task performance as well as the degree to which participants understand the knowledge they gain through the gamification process.

Thirdly, as data is gathered on participants the integrators need to develop a scorecard so as to rank performance eby individual/groups and further ascertain the degree of reinforcement as well as reward associated with same. Moreover, it is integral to the scorecard process that integrators assess prior reward criteria and revise accordingly in anticipation of new performance expectation. This ensures that the degree of motivation associated with the gamification process does not become monotonous. Once the scorecard has been effectively developed they must now take time to reflect on the data gathered and its results to ascertain impact and requisite retooling/re-engineering for the next instance of engagement or intervention.

Integrators need to bear in mind that the model describes is a developmental in nature and as such subject to modifications as it evolves. Given the momentum of the gamification movement a revision of this model is inherent. However, despite the evolutionary nature of the concept the model can ascertain that adoption of each stage will provide an effective starting point to gauge the benefits of a gamified universe within education and training.

3.13   Conclusion and Final Thoughts

The opportunity to explore gamification from a conceptual measurement framework is essential to foster a better understanding of this phenomenon and allows for the development of a framework for assessment. While no empirical testing of these proposed measures articulated have yet been undertaken within a gamification context to effectively validate the measures proposed, the author believes that the approaches put forward bear validity given that they can all build on and integrate the underlying theories of Kirkpatrick, Kolb and Sweller in the overall assessment process. In fact an ideal situation will be to incorporate the work of these seminal theorists in the GPAR review to add a much richer evaluation audit that incorporates reflection, cognition and knowledge transference. Furthermore, the context of assessing/measuring the performance of gamification utilizing such a robust measurement process (i.e. Play Assessment, GPAI, Pre/Post Assessment & GPAR) allows for a 360° feedback approach to measurement which covers gamification at the player, knowledge, behavioural and strategic level, a much needed approach for validation. The approach is further supported by codification and multiple observational foci to limit bias and provide multiple views as to the potential outcomes of measurement through robust statistical analysis.

What we know thus far is that gamification is an emerging and rapidly growing trend. It may have in its short existence become an overused perhaps even abused term which makes some naysayers cringe at the thought of incentivized

learning with badges, rewards and other trinkets to support implementation. The undeniable fact is, we live in a digital age, with digital natives whose perception of learning far differs from those of us who may have been recipients of Kolb, Kirkpatrick and Sweller's theoretical assessments in traditional brick based instructional environments. The reality is those days are long behind us and unlikely to return. As learning professionals we owe it to ourselves and our emergent class of digital natives to be receptive to new methodologies of learner engagement while sticking to our philosophy of ensuring the robustness and validity of the desired learning outcomes.

References

Balakrishnan, M., Bengasamy, S., Aman, M. (2011). Teaching games for understanding in physical education: A theoretical framework and implication. Jurnal Atikan.

Braithwaite, B., & Schreiber, I. (2008). *Challenges for game designers*. Boston: Charles River Media.

Caillois, R. (1961). *Man, play and games*. New York: Free Press of Glencoe Inc.

Carroll, J. (1982). The adventure of getting to know a computer. *Computer, 15*, 49–58.

Chandler, P., & Sweller, J. (1991). Cognitive load and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Choi, I., & Lee, K. (2009). Designing and implementing case-based learning environments fror enhancing ill-strutured problem solving: Classroom management probles for prospective teachers. *Education Tech Research Dev, 57*, 99–129.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper Perrennial.

DeLeeuw, K., & Mayer, R. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology, 100*, 223–234.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification". *MindTrek'11 Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9-15). New York: ACM Digital Library.

Hanari, J., Koivisto, J., Sarsa, H. (2014). Does gamification work?—A literature review of empirical studies on gamification. *47th Hawaii International Conference on Systems Science.* Hawaii, USA.

Hirsh-Pasek, K., Golinkoff, M., Berk, E., Singer, G. (2009). A mandate for playful learning in school: Presenting the evidence. New York, NY. Oxford University Press.

Huotari, K., Hamari, J. (2012a). Defining gamification: A service marketing perspective. *MindTrek'12 Proceeding of the 16th International Academic MindTrek Conference* (pp. 17-22). New York: ACM Digital Library.

Huotari, K., Hamari, J. (2012b). "Defining Gamification: a service marketing perspective." *16th International Academic MindTrek Conference* (pp. 17-22). Tampere, Finland: ACM.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning & development*. Englewood Cliffs, NJ: Prentice-Hall.

Lewis, K., Lange, D., & Gillis, L. (2005). Transactive memory systems, learning, and learning transfer. *Organizational Science, 16*, 581–598.

Mayer, I. (2012). Towards a comprehensive methodology for research and evaluation of serious games. *Procedia Computer Science* (pp. 233-247). Elsevier.

Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), Cambridge handbook of multimedia learning (pp. 31–48). New York: Cambridge University Press.

McCarthy, M. (2010). Experiential learning theory: from theory to practice. Journal of Business & Economic Research, Vol 8 No. 5.

Memmert, D., & Harvey, S. (2008). The game performance assessment instrument 9GPAI): Some concerns and solutions for further development. *Journal of Teaching in Physical Education, 17*, 220–240.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance. *Educational Researcher, 23*, 13–23.

Metzler, M. (2000). Instructional models for physical education. Boston, MA: Allyn and Bacon.

Mislevy, J., & Haertel, D. (2006). Implications for evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*, 6–20.

Mislevy, J., Steinberg, S., & Almond, G. (2003). On the structure of educational assessment. *Measurement: Interdisciplinary Research and Perspective, 1*, 3–62.

O'Neil, H., Baker, E., & Wainessa, R. (2005). Classification of learning outcomes: evidence from the computer games literature. *The Curriculum Journal., 16*.

Oblinger, D. (2004). The next generation of educational engagement. *Journal of Interactive Media in Education., 8*.

Rahimic, Z., Vuk, S. (2012). Evaluating the effects of employee education in B&H companies. *Conference Proceedings: International Conference of the Faculty of Economics Sarajevo (ICES)* (pp. 1044-1057). Sarajevo: University of Sarajevo, Faculty of Economics.

Shute, V. (2011). Computer games & publishing. Charlotte, NC: Information Age.

Shute, V., Masduki, I., & Donmez, O. (2010). Conceptual framework for modeling, assessing and supporting competencies within game environments. *Cognition and Learning, 8*(2), 137–161.

Shute, V., Ventura, M., Bauer, M., Zapata-Rivera, D. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning: Flow and grow. In M. C. U. Ritterfeld, *Serious games: Mechanisms and effects* (pp. 295-321). MaHwah, NJ: Routledge.

Sweller, J. (1999). Instructional design in technical areas. Camberwell, Victoria, Australia: Australian Council for Educational Research.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), Cambridge handbook of multimedia learning (pp. 19–30). New York: Cambridge University Press.

Tan, K., & Newman, E. (2013). The evaluation of sales force training in retail organizations: A test of Kirkpatrick's Four-level Model. *International Journal of Management, 30*.

Weisberg, D., Hirsh-Pasek, K., & Golinkoff, R. (2013). Embracing complexity: Rethinking the relation between play and. *Psychological Bulletin, 139*, 35–39.