

Stylized Facts for Mobile Game Analytics

Anders Drachen^{*}, Nicholas Ross[†], Julian Runge[‡] and Rafet Sifa[§]

^{*}Aalborg University

Email: andersdrachen@gmail.com

[†]University of San Francisco, San Francisco, California 94105

Email: ncross@usfca.edu

[‡]Wooga, Germany

Email: julian.runge@gmail.com

[§] Fraunhofer IAIS

Email: rafet.sifa@iais.fraunhofer.de

Abstract—There are numerous widely disseminated beliefs in the rapidly growing domain of Mobile Game Analytics, notably within the context of the Free-to-Play model. However, the field remains in its infancy, as there is limited conclusive empirical knowledge available across industry and academia, to provide evidence for these beliefs. Additionally, the current knowledge base is highly fragmented. For Mobile Game Analytics to mature, empirical frameworks are needed. In this paper the concept of stylized facts is presented as a means to develop an initial framework for a common understanding of key hypotheses and concepts in the field, as well as organizing the available empirical knowledge. A focus on stylized facts research will not only facilitate communication but also, more importantly, improve the quality and actionability of insights. Unified terminology and a comprehensive collection of stylized facts can be the building blocks for a conceptually well-founded understanding of mobile gaming.

I. INTRODUCTION

Game Analytics as a domain of research and inquiry has rapidly emerged within the past ten years, going from being virtually unknown in industry and academia to forming a core part of game development and research [1]–[8]. Within the larger umbrella of Game Analytics, mobile games have grown to form a substantial part of the industry in terms of both revenue and number of games developed. For example, in 2015 the Apple App Store carried almost 400,000 games on its mobile platform, a platform that did not exist ten years prior [9]. Academic and industry games research has expanded in parallel [10]–[13]. The vast majority of mobile games follow the freemium business model, i.e. they are Free-to-Play (F2P,FtP) and generate revenue via In-App Purchases (IAPs) [14]. F2P games are thus dependent on the willingness of users to make purchases, and therefore require analytical support. This has prompted a surge in research on player behavior, with a focus on profiling, monetization, funnel analysis, onboarding, prediction and the impact of design [1], [5], [8], [15].

A byproduct of this rapid development is a general immaturity in the field, the common symptom of which is an uneven level of analytical capacity across both industry and academics. This is especially true for Small-Medium sized Enterprises (“SMEs”) which have trouble keeping up due to the specialized knowledge and investment required

to take advantage of behavioral data. On the academic side, researchers struggle to keep up with the work being done in a field where industry has access to both more resources and more data. In essence, the field is in its infancy and the available knowledge is heavily fragmented, not the least due to Game Analytics interdisciplinary nature and the lack of knowledge sharing between academia and industry. One result of this immaturity is a lack of standardization and effective communication between industry and academics [8].

The root cause of these problems is the lack of a framework for organizing current knowledge and prioritizing interesting problems. This is unlike older and more mature fields such as Economics, where the research agenda is better defined and open questions more broadly known. For example, in the field of economic growth, specific questions regarding why poor countries are poor and rich countries are rich are well defined and deemed important, allowing researchers to prioritize [16]. Similar, in Game AI there are specific, well-known flagship areas based on previous research [17]. Researchers in those fields, given a hypothetically perfect data set, would use that data to solve these questions first. In the field of Mobile Game Analytics, this prioritization is not explicit.

The current situation is common for new empirical research domains, but there are clear benefits to establishing research frameworks for guiding the development of new research and organizing existing knowledge. In the case of Mobile Game Analytics, where no such framework exists, building one involves testing prevalent ideas about player behavior which currently have limited empirical backing.

In this paper the current state-of-the-art of Mobile Game Analytics is described in terms of the distribution of knowledge in the domain, the general forces driving current research and interest in mobile games. The concept of *stylized facts* is introduced in the context of Mobile Game Analytics for the purpose of providing a vehicle for developing a framework, across both industry and academics, for research in the area. To be specific, such a framework should be an aggregation of current knowledge into recognizable areas, a synchronization of terminology and definitions, and a system for organizing and defining the open problems in the domain.

Unfortunately, given the fragmented state of empirical knowledge in Mobile Game Analytics, it is not currently possible to define stylized facts. To facilitate their creation, this study introduces two proto-stylized fact concepts which describe situations where less empirical validation is available: *beliefs* and *hypothetical stylized facts*. These, respectively, represent situations where none/highly limited vs. some empirical support is available. Namely, they represent lower levels of empirical validation, which are needed given the current state of available empirical evidence in Mobile Game Analytics. Introducing these concepts should provide a roadmap for structuring current knowledge and building towards a situation where stylized facts can be generated and validated.

Stylized facts originate in the Social Sciences, notably Economics [18], [19]. [18] defined stylized facts as: “*such properties, common across a wide range of instruments, markets and time periods are called stylized empirical facts.*” They are observations that have been made in so many contexts that they are widely understood to be empirical truths, and used as the basis for other theories. Stylized facts are essentially a simplified presentation of a broad, empirical finding, similar to basic theories describing the relationship between variables [20]. They are sometimes referred to as statistical regularities and can provide a basis for aggregating and communicating knowledge in a field, being a particularly popular tool in Economics [18], [19]. Collections of stylized facts form frameworks of organized knowledge in Economics and help define areas of open research. They thus bear similarity to the use of standards in the domain of Human-Computer Interaction and Computer Science, although less specifically defined [21].

The Game Analytics field, not only in mobile/F2P games but perhaps especially therein, is rife with empirical ideas and statements that are presented like stylized facts, but have yet to be rigorously validated. For example, in F2P games, it is a widely held belief that player retention correlates with monetization, i.e. as the frequency users play a game increases so does the revenue generated by that game. If this belief was based on a substantial amount of empirical research, even if the details varied slightly over different contexts (e.g. match-3 games vs. collectible card games, PC versus mobile platforms), it could be labeled as a stylized fact. However, there is currently little available empirical evidence in mobile games to convincingly define this as a stylized fact.

A framework, such as that provided by stylized facts, is needed to mature research in this domain. Towards this end we propose, based on current evidence and industry interest, a set of hypothetical stylized facts that should help focus research on open problems. Studying these problems will push research towards turning these hypothetical stylized facts into actual stylized facts and thus establish a basic organizational framework. Luckily, there is already a base amount of empirical research done in the areas that we consider ripe for the establishment of stylized facts; our study builds upon that

effort.

The organization of this paper is as follows. Section 2 provides background and a literature review of Mobile Games Analytics, Section 3 defines stylized facts, with a focus on showing how they help mature a field, while Section 4 presents a set of hypothetical stylized facts of interest to both academic and industry researchers. Finally, Sections 5 and 6 provides a vision for future research and a conclusion, respectively.

II. BACKGROUND: THE STATE OF MOBILE GAME ANALYTICS

There are numerous reasons for the interest in mobile games research, most of which stem from the unique opportunity they present for studying behavior, including:

- The technological underpinnings of the mobile platform and its dissemination into society, coupled with the introduction and spread of games on these platforms, provides the opportunity for behavioral data collection at an unprecedented scale.
- Mobile games form a unique opportunity to inform game design and game research: devices are ubiquitous, both across geography and culture, and are evolving very rapidly with new devices being introduced in the marketplace continually.
- Mobile devices are also often continually network connected, which enables constant behavioral measurement.
- Mobile games are generally cheaper to develop when compared to consoles and, thanks to pre-installed distribution services such as the Apple App Store, mobile games are easy to find (if the player knows the name of the game - there are hundreds of thousands of mobile games available making exploratory searches difficult), install and uninstall. The games are also, in the majority of cases, F2P, implying that the barrier to entry for a user is only the time it takes to download [14]. This has yielded a situation where a user’s investment in a title is low, impacting the design of mobile games and the approach to revenue generation on these platforms (commonly referred to as monetization) [5], [6], [8].

Jointly, these aspects of mobile games provide a unique opportunity for large-scale research and, thanks to the technology powering mobile devices, detailed tracking of user behavior. In other words, games researchers have access to broader and deeper datasets concerning user behavior than ever before; and this across audiences that are, while already large and diverse, continually growing.

However, due to the lack of standardized knowledge and shared frameworks in Mobile Game Analytics, a primary challenge facing all games researchers is deciding where their energy should be spent, i.e. how to utilize this available data, or discover what data is needed to address a given question, or even what the most important open questions are [5], [6], [8], [22], [23]. Even basic best practices for using the currently available, overwhelming amount of behavioral telemetry data for games is lacking [1], [8], [13]. This is notably critical given the diversity of games and contexts in which they are played.

On the industry side, the rise of analytics has led to the hiring of data analysts, machine learning specialists and data mining experts at prodigious rates to take advantage of these new data sources. To be successful in the F2P space requires actively leveraging data to inform decision-making [5]. Today large publishers have dozens of data scientists and user researchers working for them, far outstripping the academic ranks. However, SMEs are struggling to catch up, especially small developers operating on constrained budgets [24].

Despite the surge of interest in Mobile Game Analytics, there remains a dearth of available information from the industry. This can, to a large degree, be explained by the business value of behavioral data from games, the sensitivity of personal data, and by the relative recent introduction of analytics in games. As yet, there are no associations of game analytics or other formal bodies that can help promote knowledge sharing or standardization. Confidentiality around managerially sensitive metrics such as revenue and churn/retention makes knowledge sharing difficult apart from high-level discussions, as evident in presentations at industry events, white papers, reports and blog posts released by analysts, e.g. [25], [26]. Similarly, aggregate information on player behavior across many games is only available for publishers, or via analytics companies and generally locked behind paywalls.

In summary, industry knowledge has a tendency to stay within the confines of developers or publishers rather than be disseminated broadly. However, even when efforts are made to communicate information, the lack of shared definitions, terminology and understanding are obvious. The lack of consistent terminology signifies that we are not operating from the same baseline because of the lack of underpinning stylized facts. This leads to a continual loop of companies and researchers re-inventing the same concepts, terms and solutions.

Academic Mobile Game Analytics research finds itself in a generally similar situation. This includes the level of secrecy, as the majority of academic papers contributing to the field of Game Analytics (including F2P games), do not release the datasets used. This is commonly because the research is carried out in collaboration with a company which needs to keep the raw data confidential, or because additional studies can be performed and published on the dataset. Furthermore, similar to industry, the interdisciplinary nature of Mobile Game Analytics means that publications are distributed across numerous databases, journals, publishers and indexes and thus challenging to discover. Academic research is also often locked behind publisher paywalls, which means that the knowledge generated is not readily available to the industry, especially SMEs. It is also the case that academic research in Mobile Game Analytics tends to redefine key concepts such as churn, retention, life-time value etc. in each new paper, rather than coalescing towards shared definitions.

Academic research work in Mobile Game Analytics, F2P games included, is currently fragmented, covering a wide variety of business intelligence problems, rather than aligned

around a set of open problems. This is to be expected in the explorative phase of a new domain being established, but means that even when studies overlap they tend to define key terms differently. Examples from interdisciplinary studies include behavioral profiling [1], [27], [28], player activity analysis [29], [30], social network analysis [31] churn and retention analysis [10], [12], [32]–[34], premium user identification [13], [35], automatic game content generation [36], [37], abusive content analysis [38], opponent difficulty adjustment [39] and recommender systems [40]. It is important to note that, with the recent popularity of freemium models, more and more studies are devoted to study player behavior in F2P games [10], [12], [13], [33], [35], [38], [39]. A general observation across these studies is that they introduce individual problem definitions, and report data and methods differently, indicating a lack of standards for reporting work in Mobile Game Analytics.

III. INTRODUCING STYLIZED FACTS

As mentioned in the introduction, this study's purpose is to introduce the concept of stylized facts, as they are used in Social Science and Economics, into the context of Game Analytics. This is in order to provide a framework for structuring knowledge in the field and define open problems.

In Economics, stylized facts were introduced based on the need to create a set of common beliefs for organizing knowledge. [41] defined stylized facts as succinct encapsulations of statistical information regarding a topic: "*Stylized facts can be a useful way of organizing one's thinking about phenomena of interest, giving a broad direction to theorizing and mapping out an agenda for empirical work.*" Among the first clear uses of stylized facts was in the Economics subfield of economic growth. Explicitly writing down a list of stylized facts helped define where the field stood as well as areas for future research [42]. In particular, Kaldor [43] defined six facts regarding economic growth that any model of growth should explain. While the facts were empirical in nature, they were not perfect since, as Kaldor [43] admitted, the goal was to: "*concentrate on broad tendencies, ignoring individual detail.*" Choosing to focus on broad features allowed for a level of abstraction that pushed the field forward. Similarly, creating a set of stylized facts for the mobile video game industry has the potential to push forward video game research in a number of different ways:

1) **Research evaluation:** Having a framework to evaluate research creates a well-defined environment for valuing incremental contributions. In particular, research is often evaluated using a Bayesian framework where value is ascribed to how the result changes our beliefs [44]. Research that either strongly confirms a *weak prior belief* or cast doubts upon *strong prior beliefs* is valued above research that either confirms an already strong prior or weakly counteracts a weak prior. When researchers have a set of stylized facts to base their research on they can be leveraged as priors, allowing us to better understand the contribution of a piece of research. Specifically, not having a set of stylized facts creates

an environment where research tends toward exploration and technique, rather than interpretation.

To take an example, [12] provided the first formal definition of the churn problem in games and provided prediction models across five mobile game titles. As part of the contribution, the authors define a set of behavioral features and describe their influence on the process of classifying churning players, i.e. for predicting future player departure. Despite the importance of understanding player churn, particularly for practitioners, and the existence of other recent papers in this area [10], [13], [33], it is difficult for non-experts to evaluate the importance of the work because there is no baseline to evaluate it against. Furthering this example, consider the case of applying Hidden Markov Models to churn prediction by [10]. Does this research provide a new understanding of player behavior in F2P mobile games, or is the contribution a confirmation of something already known but in a new context? Trying to appraise research without defined prior stylized facts (or another system for organizing and evaluating empirical knowledge), forces us to focus on the technique used by the authors rather than the interpretation of the results of that technique. In other words, an author can present a novel method for solving a problem of interest, but without a set of stylized facts, evaluating the method's contribution can be challenging.

2) **Research focus:** Having a framework of stylized facts allows researchers - whether in academia or industry - to focus their research energy. For example, if we accept the idea that “retention in core games is lower than in more casual titles,” then studies confirming, denying or defining when this fact is true become priorities for researchers (when given a dataset that permits investigating this broad fact). In much the same way that the equity premium puzzle [45] provides a prioritization mechanism for research in the Finance field, so would the establishment of stylized facts in Game Analytics.

3) **Standards and reporting:** Having stylized facts provides a consistent framework for information to provide regarding a game title when reporting research. In many other applied fields, studies attempt to present consistent, basic information about the subjects in their study. For example, empirical studies focusing on mergers and acquisition activities report the Herfindahl–Hirschman index for that industry, as it provides a numerical data point addressing the current industry concentration. While researchers in this field know that this number is not a perfect encapsulation of industry concentration, it does provide baseline information that is useful for other researchers.

In games research, the lack of stylized empirical facts has hindered the presentation of this type of foundational information about the subject of research.

The natural question arising from the above considerations is how to identify and define stylized facts in mobile games. Fortunately, Game Analytics is an applied field in an area where analytics are ubiquitous: simply putting a game up on the Apple App store or the Google Play store generates basic KPIs about a title such as downloads and revenue. Many applications also contain more advanced tracking systems [46]

and thus game developers are generally familiar with the concept of analytics and understand the importance of it, even if analytics is not deeply integrated within every company [5], [6], [8].

Stylized facts rest on a substantial body of empirical knowledge which amalgamates to high-level, crisp, facts. The current state-of-the-art in Game Analytics for mobile/F2P games is not in a place where enough empirical work has been done to support stylized facts as they occur in Economics. If we were to represent stylized facts in games currently, they would sound like facts, but would have a much weaker empirical basis than in other fields. In essence, while we cannot currently present a series of stylized facts for Game Analytics, we can provide examples of what they would look like and what they could be used for and include any limited empirical evidence that is currently available. Doing so provides a starting point or guide for researchers to either validate (or invalidate) them.

We will therefore refer to these as *hypothetical stylized facts*. Where stylized facts are normally defined bottom-up, these are defined top-down, i.e. based on current high-level ideas and perceptions in the F2P Game Analytics domain. While there maybe some available empirical evidence supporting these hypothetical stylized facts, it is clearly not enough to rigorously, generally, support them. These hypothetical stylized facts will need to be examined in the future, and refuted or confirmed for different contexts. They will change and some may gradually mature into actual stylized facts.

An example of a hypothetical stylized fact is represented in [1], who examined the belief that playtime distributions follow a power law. The authors indicated prior work across academia and industry that described or analyzed playtime distributions. They then demonstrate across over 3,000 games that playtime distribution could be modeled using a Weibull distribution. This led the authors to propose a “playtime principle” suggesting that playtime in games follows a Weibull distribution, and furthermore described the potential cause in terms of rising and falling components of human interest. Importantly, the authors acknowledge that this principle needed further validation. This is an example of researchers taking a prevalent but minimally supported belief, compiling what empirical evidence exists and adding to it, and using this as a the basis for proposing a hypothetical stylized fact. Other potential stylized facts in Game Analytics exist in behavioral profiling [1], [47] and churn prediction [10], [12], [13] as well as in network balancing in Massively Multi-Player Online Games (MMOGs) [48].

We contrast hypothetical stylized facts against *beliefs*. Beliefs are, in this context, perceptions with little or no documented empirical evidence. There are many beliefs in Mobile Game Analytics that often get presented as stylized facts without the required empirical basis. These include statements such as:

- “Hard core games have higher monetization” [49]
- “Monetization and retention move in different directions” [50]
- “Casual Users are not as engaged as other users” [51]

- “Tablet Users Monetize better than SmartPhone Users” [52]

Beliefs can give rise to hypothetical stylized facts through amalgamation and analysis of empirical knowledge, which can then, through further research, become stylized facts.

A. The importance of context

Whereas stylized facts in Economics are often presented as being context-independent, this is only partially correct and commonly pointed out in the literature. The idea is not that they apply everywhere but are a general trend, and that exceptions to the trend are of interest [42]. In Mobile Game Analytics, beliefs are often presented as stylized facts that are context-independent, e.g. the higher the difficulty of a game, the quicker the drop in retention. However, games are highly varied in their design and how they approach building user experience. Games can also be played alone or with others, physically or online. While research on the effect of context on gameplay and user experience remains somewhat limited [53], context would appear to play a significant role in player behavior. The expectation is that stylized facts in Mobile Game Analytics need to be accompanied by definitions for these different contexts or conditions, or even for some stylized facts to be defined directly in relation to a specific set of conditions.

IV. HYPOTHETICAL STYLIZED FACTS: A FIRST ATTEMPT

In this section, we create a set of hypothetical stylized facts that we believe, if studied and researched, mature the field of Mobile Game Analytics. In particular, our stylized facts relate monetization, retention and engagement, which are frequently studied academically while also being of prime interest in industry. Despite their importance, however, there are opposing views on how they relate to each other [50], [54] and it is apparent that a comprehensive answer relating all three is not straightforward [55].

There is already some literature present around this topic. Most analytical models of free-to-play build on the assumption that retention comes first and determines monetization [56]. However, a look into the marketing literature reveals that the causality between retention and monetization is unlikely to be unidirectional [57]–[59]. A number of papers in this field also relate to how pricing and sunk costs (such as an initial purchase or previous engagement) affect the level and pattern of consumption [57], [60]. Paywalls, which block content from customers in much the same way that some F2P games monetize, have also been studied in the literature [61], [62] while industry sources have more directly considered the effect of blocking core gameplay through monetization practices [63]. In other words, while there is some literature discussing the relationship between retention and monetization, it has yet to be clearly defined.

Further, engagement (as often measured by sessions, rounds or time spent in game per day) and retention are generally assumed to go hand in hand. [50] While this may largely hold, there are clear instances where a game with lower retention is characterized by more engagement. As an example, consider

the iOS versions of Wooga’s Pearl’s Peril and Jelly Splash. While Jelly Splash has a more than ten percentage points higher day one retention than Pearl’s Peril, players of Pearl’s Peril play more than twice as many sessions per day as Jelly Splash players.

Given the incomplete and fragmented empirical information regarding these important topics, we propose the following analytical notation to facilitate the creation of stylized facts. Let monetization be defined as μ , retention as ρ and engagement as γ . Revenue (and success) of a free-to-play game can then be written as $\pi = f(\rho, \mu, \gamma)$. Using this notation, we propose the following stylized facts, writing them as simple derivatives:

- $\frac{d\pi}{d\rho} > 0, \frac{d\pi}{d\gamma} > 0, \frac{d\pi}{d\mu} > 0$: Revenue π derived from a free-to-play mobile game is increasing in retention ρ , in engagement γ and in monetization μ [56], [64].
- $\frac{d\mu}{d\rho} > 0, \frac{d\mu}{d\gamma} > 0$: Monetization (potential) μ is increasing in both retention ρ and engagement γ [56], [64].
- $\frac{d\rho}{d\mu} \geq 0, \frac{d\gamma}{d\mu} \geq 0$: Retention ρ and engagement γ are non-decreasing in changes to monetization μ (abstracting from dubious monetization mechanisms that may adversely affect usage).
- $\frac{d\rho}{d\gamma} \geq < 0, \frac{d\gamma}{d\rho} \geq < 0$: Retention ρ and engagement γ can increase or decrease in each other or even be stable contingent on different levels of each other.

Functions and derivatives provide a concise, analytically rigorous, notion to define the associations between these important levers that can be easily simplified when presenting to lay audiences. While this is a condensed example, it provides the basis for further discussion about what kind of format stylized facts in Mobile Game Analytics should take and how they could be published to the community.

In the above sections, we have provided citations to published information where possible, but a number of the hypothetical stylized facts discussed above are known in the industry, but, because of confidentiality concerns, are more difficult to find direct information regarding. Later we will talk more about these types of informational asymmetry issues.

V. DISCUSSION: A VISION FOR THE FUTURE

In order to push our field forward, we need to move beliefs to hypothetical facts towards stylized facts. This will require collaboration between academia and industry across a number of venues of inquiry, from improving collaborations and fostering data sharing, jointly shaping definitions and building the framework for standardization.

The previous section laid out four hypothetical stylized facts, but in doing so, it obfuscated an important issue currently surrounding Mobile Game Analytics, the lack of consistent definitions. As in other new and immature fields, one of the major hurdles to achieving a set of stylized facts is the lack of consistent definitions. Consider the case of retention, or the likelihood that user returns to a title. Depending on who you ask, you get a different definition [65]: “*The thing is that, however paramount it is for app publishers, retention has a problem. Everyone talks about it, but there seems to be no clear consensus on a common definition of what retention*

really means nor how it is actually calculated. The truth is that there are several ways to compute this metric, all of which lead to sensibly different results. In turn, people often end up comparing apples with oranges.”

Other practitioners have also lamented the inconsistencies surrounding the definition of retention, monetization and engagement. Given that these terms are used throughout both industry and academics, the fact that there is not an easily understood definition is a reminder of the research opportunity in this field. As researchers we can help define these measures by studying how different possible definitions correlate with each other and other business objectives. Importantly, researchers can also pursue understanding where and when different definitions break down or when they should not be used.

This type of foundational work is common in other applied fields. For example, looking at the history of gross domestic product reveals a litany of discussions regarding how different types of production should be measured [66], let alone how to interpret and use it. For video games, similar discussions regarding monetization, engagement and retention would create a common knowledge base that would push forward both academic and industry research. However, achieving these definitions require collaboration across academia and industry.

The opportunities in mobile games across industry and academia are substantial, as are the challenges. These are not issues that can be dealt with rapidly but rather require a shift in operational and strategic mindset. In particular, the primary reason for the immaturity of Mobile Game Analytics is a weakness in industry-academic partnerships.

In fact, industry usually has earlier and better access to the most recent developments and information and thus tends to be ahead of academic research. The raw capacity for analytics work in the industry massively outstrips that in the associated academic areas. Industry research, on the other hand, usually lacks the analytical structure and rigor present in more academically focused research. To increase the usefulness of academic research, the goal should be to leverage industry's experience to find facts that are in the intersection of what researchers can do and what industry finds useful.

In our experience, many of the industry-academic partnerships are closer to a “hand-off” where academics are given data, with some minor restrictions and allowed to “do research on it,” with practitioners exerting little influence past this point, or expecting much of a return. While this type of relationship can yield some interesting techniques and insight, the lack of an ongoing cooperation between academic and industry researchers has a negative effect on both groups. On the academic side, the one-off, almost transactional, nature of these relationships leaves researchers in a bind: while they would like to engage in research that would push forward academic and industry agendas, because they are not involved on the industry side they are left to their own devices, attempting to “guess” what industry would like to see. On the other hand, industry practitioners are not finding academic research that useful, as long academic time lines and the directionless nature

of the relationship ends up influencing the type of research that is being done.

Industry and the academic community need to work together more efficiently to take the understanding of mobile games to the next level. To make our proposition actionable, it is of paramount importance to point at tools that enable the joint development of stylized facts. While stylized facts are everywhere - they appear in many professional conversations, often in the form of intuitive assumptions - generating and collecting them in a structured manner is a challenge. We put forward two suggestions for action in this regard:

1) **Improved data provisioning from companies:** While much work has been done in Game Analytics, many gaming companies do not yet understand the strategic relevance of making their data accessible to researchers and allowing them to write about the insights gleaned. Similarly, academics may not see the relevance of working with industry. It should be especially noted that relating findings to basic game characteristics is important to encourage scientific debate. For instance, studies that fail to include core information regarding a game, such as genre or basic statistics regarding retention and monetization lower their contribution potential. In particular, academic researchers should make, as a condition for doing research, the publication of basic statistical information about a title (any personal information should be anonymized); and industry should support this requirement in the interest of maturing the domain as a whole. On the research evaluation side, journals and conferences should encourage authors to include this information. Studies without this basic information is of less value to both practitioners and other academic researchers.

2) **Research cooperations:** Research cooperations, where academics work with a company, are becoming more widespread. However, like other industry-academic partnerships, they sometimes suffer from a lack of communication (and particularly communication depth) to ensure an aligned and deep understanding of data and insights. In order to yield a successful research cooperation, two key features are desirable. First, academic researchers have to be sympathetic to the realities of industry. This means focusing on research efforts that maybe more short-term and less rigorous, with the understanding that the academic researcher can revisit a topic on their own when deciding to pursue publication. Secondly, industry partners need to be sympathetic to the realities of academics. Primarily, this means assisting researchers in becoming informed about the processes and data of a company. This may mean spending time with the researcher on multiple occasions, and putting them into contact with other members of a game development team, such as data analysts, product manager and producers. If an academic researcher does not have context for the game or situation, the result is likely to be less useful. In this regard, we especially want to mention the concept of “researcher in residence where a researcher joins a company and works on-site. The physical presence of the researcher is of inconceivable value as it allows for a plethora of informal touch points within an organization. This exposes

the researcher to the challenges faced by a company in a way not possible as part of a more traditional research cooperation, and - potentially more importantly - facilitates the delivery of valuable impulses derived from thorough research.

These suggestions are complimentary to the current set of activities that assist in the creation of stylized facts. On the industry side, conferences, such as the Game Developers Conference and analytics-focused talks and panels increase connectivity between like-minded industry-based analytics researchers. However, due to the confidentiality of sensitive information such as companies' revenue and churn/retention rates, this type of data is neither consistent or plentiful.

Putting together fruitful industry-academic partnerships is one first step toward creating precise and useful definitions. If academics attempt to define these key terms in isolation, the resulting terms may not conform to how industry uses them, further pushing academics and practitioners apart. For Mobile Game Analytics to mature as a field, there must be an increase in industry-academics partnerships to, first, define key terms and secondly, turn beliefs into stylized facts. Without an increase in this type of partnership, Mobile Game Analytics researchers will fail to take advantage of the opportunity presented to them.

VI. CONCLUSION

Mobile Game Analytics has emerged rapidly within the past decade, from virtually unknown to playing a foundational role in what has become a major part of the game sector in both industry and academia. Given the formative stage of the domain, it is not surprising that current knowledge is fragmented, often not publicly accessible, and that there is a lack of standardization. However, any domain of inquiry needs to mature, and this is also the case for Mobile Game Analytics.

In this paper, the current status of Mobile Game Analytics has been described and the challenges discussed. Furthermore, a high-level vision has been put forth for moving the domain of Mobile Game Analytics to a more comprehensive, firm base, via adopting the idea and concept of *stylized facts* and blending them into analytical models.

Adopting such a framework will enable academia, industry and especially their collaborations, to contribute to a better understanding of mobile/Free-to-Play games in a more effective manner. Without a firm definition of the concepts and terms used in Mobile Game Analytics research, industry and academic practitioners, as well as non-experts, will continue to be challenged to obtain clear information on the state of the art of knowledge in the domain or even best practices.

The proposed approach can provide both framework and direction for future research efforts. To get to that point, we, as researchers, need to spend our energy working on clearly defining terms such as monetization, retention and engagement. We also need to focus on assessing, evaluating and building empirical evidence towards constructing hypothetical and then actual stylized facts; testing and pushing to understand more within the unique opportunity that mobile games provide across academia and industry.

REFERENCES

- [1] C. Bauckhage, A. Drachen, and R. Sifa, "Clustering game behavior data," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 3, pp. 266–278, 2015.
- [2] J. Bohannon, "Game-miners Grapple with Massive Data," *Science*, vol. 330, no. 6000, pp. 30–31, 2010.
- [3] Swrve, "Swrve Monetization Report 2014," "<http://landingpage.swrve.com/rs/swrve/images/swrve-monetization-report-0114.pdf>", 2014.
- [4] G. Yannakakis, "Game AI Revisited," in *Proc. of ACM Computing Frontiers Conference*, 2012, pp. 285–292.
- [5] W. Luton, *Free-to-Play: Making Money From Games You Give Away*. New Riders, 2013.
- [6] T. Fields and B. Cotton, *Social Game Design: Monetization Methods and Mechanics*. Morgan Kaufmann, 2011.
- [7] A. Drachen, C. Thureau, J. Togelius, G. Yannakakis, and C. Bauckhage, "Game Data Mining," in *Game Analytics: Maximizing the Value of Player Data*, M. S. El-Nasr, A. Drachen, and A. Canossa, Eds. Springer, 2013.
- [8] M. S. El-Nasr, A. Drachen, and A. Canossa, *Game analytics: Maximizing the value of player data*. Springer Science & Business Media, 2013.
- [9] Statista, "Number of available apps in the itunes app store from 2008 to 2015 (cumulative)," "<http://www.statista.com/statistics/268251/number-of-apps-in-the-itunes-app-store-since-2008/>", July 2015.
- [10] J. Runge, P. Gao, F. Garcin, and B. Faltings, "Churn Prediction for High-value Players in Casual Social Games," in *Proc. of IEEE CIG*, 2014.
- [11] R. Sifa, C. Bauckhage, and A. Drachen, "The Playtime Principle: Large-scale Cross-games Interest Modeling," in *Proc. of IEEE CIG*, 2014, pp. 365–373.
- [12] F. Hadiji, R. Sifa, A. Drachen, C. Thureau, K. Kersting, and C. Bauckhage, "Predicting Player Churn in the Wild," in *Proc. of IEEE CIG*, 2014.
- [13] H. Xie, S. Devlin, D. Kudenko, and P. Cowling, "Predicting Player Disengagement and First Purchase with Event-frequency Based Data Representation," in *Proc. of CIG*, 2015.
- [14] Y. LeJacq, "Something for nothing: How the videogame industry is adapting to a 'freemium' world," *International Business Times*, September 2012.
- [15] E. Seufert, *Freemium Economics: Leveraging Analytics and User Segmentation to Drive Revenue*. Morgan Kaufmann, 2014.
- [16] P. J. Klenow and A. Rodriguez-Clare, "Economic growth: A review essay," *Journal of monetary economics*, vol. 40, no. 3, pp. 597–617, 1997.
- [17] G. N. Yannakakis and J. Togelius, "A panorama of artificial and computational intelligence in games," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 4, pp. 317–335, 2015.
- [18] R. Cont, "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001. [Online]. Available: <http://dx.doi.org/10.1080/713665670>
- [19] M. Sewell, "Characterization of financial time series," *UCL Department of Computer Science, Research Note*, no. RN/11/01, 2011.
- [20] K. S. Borden and B. B. Abbott, *Research design and methods: A process approach*. McGraw-Hill, 2002.
- [21] A. Dix, T. Roselli, and E. Sutinen, "E-learning and human-computer interaction: Exploring design synergies for more effective learning experiences," *Educational Technology & Society*, vol. 9, no. 4, pp. 1–2, 2006.
- [22] B. Chong, "Game analytics 101," "<https://www.raywenderlich.com/7559/game-analytics-101>", January 2012.
- [23] A. Bilas, "You don't need to be a data scientist," "<http://www.gamesindustry.biz/articles/2015-06-22-you-dont-need-to-be-a-data-scientist>", June 2015.
- [24] A. Rajpurohit, "Interview: Pallas horwitz, blue shell games on why data science is so critical for gaming studios," "<http://www.kdnuggets.com/2014/08/interview-pallas-horwitz-data-science-gaming-studios.html>", Aug 2014.
- [25] L. Mellon, "Applying metrics driven development to MMO costs and risks," <http://maggotranch.com/>, 2009.
- [26] G. Zoeller, "Game Development Telemetry," in *Game Developers Conference*, 2011.

- [27] C. Bauckhage and R. Sifa, "k-Maxoids Clustering," in *Proc. of KDML-LWA*, 2015.
- [28] A. Drachen, R. Sifa, C. Bauckhage, and C. Thureau, "Guns, Swords and Data: Clustering of Player Behavior in Computer Games in the Wild," in *Proc. of IEEE CIG*, 2012.
- [29] G. Wallner, "Sequential Analysis of Player Behavior," in *Proc. of ACM CHI Play*, 2015.
- [30] A. Drachen, G. N. Yannakakis, A. Canossa, and J. Togelius, "Player Modeling using Self-Organization in Tomb Raider: Underworld," in *Proc of IEEE CIG*, 2009.
- [31] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, "Alone together?: Exploring the Social Dynamics of Massively Multiplayer Online Games," in *Proc. of ACM SIGCHI HFICS*, 2006.
- [32] E. Harpstead, T. Zimmermann, N. Nagapan, J. J. Guajardo, R. Cooper, T. Solberg, and D. Greenawalt, "What Drives People: Creating Engagement Profiles of Players from Game Log Data," in *Proc. of ACM CHI Play*, 2015.
- [33] P. Rothenbuehler, J. Runge, F. Garcin, and B. Faltings, "Hidden Markov Models for Churn Prediction," in *Proc. of SAI IntelliSys*, 2015.
- [34] R. Sifa, C. Ojeda, and C. Bauckhage, "User Churn Migration Analysis with DEDICOM," in *Proc. of ACM RecSys*, 2015.
- [35] R. Sifa, F. Hadiji, J. Runge, A. Drachen, K. Kersting, and C. Bauckhage, "Predicting Purchase Decisions in Mobile Free-to-Play Games," in *Proc. of AAAI AIIDE*, 2015.
- [36] N. Shaker and M. Abou-Zleikha, "Alone We Can Do so Little, Together We Can Do so Much: A Combinatorial Approach for Generating Game Content," in *Proc. of AAAI AIIDE*, 2014.
- [37] G. N. Yannakakis and J. Togelius, "Experience-Driven Procedural Content Generation," *IEEE Transactions on Affective Computing*, 2011.
- [38] K. Balci and A. Salah, "Automatic Classification of Player Complaints in Social Games," *IEEE Transactions on Computational Intelligence and AI in Games*, 2015.
- [39] P. I. Cowling, S. Devlin, E. J. Powley, D. Whitehouse, and J. Rollason, "Player Preference and Style in a Leading Mobile Card Game," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 3, pp. 233–242, 2015.
- [40] R. Sifa, C. Bauckhage, and A. Drachen, "Archetypal Game Recommender Systems," in *Proc. of KDML-LWA*, 2014.
- [41] P. Geroski, "What do we know about entry?" *International Journal of Industrial Organization*, vol. 13, no. 4, pp. 421 – 440, 1995, the Post-Entry Performance of Firms. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/016771879500498X>
- [42] C. I. Jones and P. M. Romer, "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," in *American Economic Association*, 2009.
- [43] N. Kaldor, *The Theory of Capital: Proceedings of a Conference held by the International Economic Association*. London: Palgrave Macmillan UK, 1961, ch. Capital Accumulation and Economic Growth, pp. 177–222.
- [44] D. Burgstahler, "Inference from empirical research," *Accounting Review*, pp. 203–214, 1987.
- [45] R. Mehra and E. C. Prescott, "The equity premium in retrospect," *Handbook of the Economics of Finance*, vol. 1, pp. 889–938, 2003.
- [46] C. Klotzbach, "Enter the matrix: App retention and engagement," ["http://flurrymobile.tumblr.com/post/144245637325/appmatrix"](http://flurrymobile.tumblr.com/post/144245637325/appmatrix), May 2016.
- [47] R. Sifa, A. Drachen, C. Bauckhage, C. Thureau, and A. Canossa, "Behavior evolution in tomb raider underworld," in *Proc. of IEEE CIG*, 2013.
- [48] J. Feng and D. Saha, "A Long-term Study of a Popular MMORPG," in *Proc. ACM SIGCOMM WNSSG*, 2007.
- [49] "How genre affects f2p game genres: A comparison," <https://deltadna.com/blog/the-best-f2p-game-genres-kpis/>, 2015.
- [50] "Forget monetisation - f2p success is all about retention and engagement," ["http://www.pocketgamer.biz/mobile-mavens/60951/forget-hahamonetisation-its-all-about-retention-and-engagement/"](http://www.pocketgamer.biz/mobile-mavens/60951/forget-hahamonetisation-its-all-about-retention-and-engagement/), March 2015.
- [51] "Casual versus core," <http://www.gamasutra.com/>, 2000.
- [52] "'exclusive: Nexus 7 monetizes better than other android tablets, says tinyco'," ["http://www.adweek.com/socialtimes/exclusive-nexus-7-monetizes-better-than-other-android-tablets-says-tinyco/535738"](http://www.adweek.com/socialtimes/exclusive-nexus-7-monetizes-better-than-other-android-tablets-says-tinyco/535738), 2013.
- [53] M. Suznjevi, L. Skorin-Kapov, and M. Matijasevic, "The Impact of User, System, and Context factors on Gaming QoE: a Case Study Involving MMORPGs," in *Proc. Netgames Conference*, 2013, pp. 1–6.
- [54] S. E. Needleman, "Mobile-game makers try to catch more whales who pay for free games," *Wall Street Journal*, May 2015.
- [55] "Quora: How do you model the relationship between engagement/retention and monetization?" <https://www.quora.com/How-do-you-model-the-relationship-between-engagement-retention-and-monetization>, 2013.
- [56] E. Seufert, "A comprehensive free-to-play game model: revenue, dau, virality, and retention," <http://mobiledevmemo.com/free-to-play-spreadsheet-revenue-model/>, January 2013.
- [57] E. Shafir and R. H. Thaler, "Invest now, drink later, spend never: On the mental accounting of delayed consumption," *Journal of economic psychology*, vol. 27, no. 5, pp. 694–712, 2006.
- [58] D. Soman, "Effects of payment mechanism on spending behavior: The role of rehearsal and immediacy of payments," *Journal of Consumer Research*, vol. 27, no. 4, pp. 460–474, 2001.
- [59] D. Soman and J. T. Gourville, "Transaction decoupling: How price bundling affects the decision to consume," *Journal of Marketing Research*, vol. 38, no. 1, pp. 30–44, 2001.
- [60] J. T. Gourville and D. Soman, "The potential downside of bundling: How packaging services can hurt consumption," *Cornell Hospitality Quarterly*, vol. 42, no. 3, p. 29, 2001.
- [61] L. Chiou and C. Tucker, "Paywalls and the demand for news," *Information Economics and Policy*, vol. 25, no. 2, pp. 61–69, 2013.
- [62] H. Oh, A. Animesh, and A. Pinsonneault, "Free vs. for a fee: The impact of information pricing strategy on the pattern and effectiveness of word-of-mouth via social media," in *Proc. of ICIS*, 2013.
- [63] "6 free-to-play design lessons from clash of clans to engage players and increase profits," <http://virtualeconomists.com/blogs/news/10823889-6-free-to-play-design-lessons-from-clash-of-clans-to-engage-players-and-increase-profits>, 12 2013.
- [64] S. Reyburn, "Exclusive: Apsalar analyzes the correlation between mobile game engagement and monetization by genre," *AdWeek*, Feb. 2013.
- [65] T. Sommer, "User retention: Yes, but which one?" ["http://www.applift.com/blog/user-retention.html"](http://www.applift.com/blog/user-retention.html), February 2014.
- [66] D. Coyle, *GDP: A brief but affectionate history*. Princeton University Press, 2015.