Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Industrial Economics and Technology Management

Lars Leegaard Marøy

Personnel attraction and selection through gamified psychometrics: A game changer for the recruitment industry?

Master's thesis in Master of Technology Management (MTM)
Supervisor: Prof. Arild Aspelund
January 2019

NTNU
Norwegian University of Science and Technology

Lars Leegaard Marøy

Personnel attraction and selection through gamified psychometrics: A game changer for the recruitment industry?

Master's thesis in Master of Technology Management (MTM)
Supervisor: Prof. Arild Aspelund
January 2019

Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Industrial Economics and Technology Management

NTNU
Norwegian University of
Science and Technology

ABSTRACT

Businesses are fighting in a "war for talent," and every single hire of a key employee is a battle of increasing strategic importance. Recruiters working in the $200 billion a year global recruitment industry help their clients in personnel attraction and selection by using a number of tools and techniques. Many of them are labour intensive and costly for the hiring organizations, who regardless have few good alternatives to paying the recruitment agencies a premium for their services.

Tech startups are now entering this scene. They offer a new take on the processes of talent attraction and identification, with products and services ranging from digitizing the steps involved in traditional hiring, to completely changing these procedures through the use of emerging technology.

Of the latter category are Pymetrics and Knack.it Corp, two US-based providers of game-based psychometric assessments for personnel attraction and selection. Their ambition is to change the rules of the recruitment game, by bringing the candidates´ *potential* to the foreground, at the expense of their *background*.

Case studies of the two services and a following focus group discussion with Norwegian industry experts were designed to answer the questions: "Which are the potential benefits and limitations of gamified psychometrics for recruitment and selection? How likely is the prospect of game-based assessment providers disrupting the recruitment industry?"

Regarding the former, the increased attraction force associated with game-based solutions over traditional non-game-based tools stands out as a significant benefit for the new tools, especially when dealing with candidates at the higher tiers of the job market. In the rest of the market, where the number of applicants can be overwhelming, the game-based tools can automate the now labour-intensive processes of designing and administrating psychometric tests, increasing their practical utility.

Out of the many possible limitations concerning the use of game-based psychometric tests, I find the *algorithmic aversion* to be one of the most substantial. Recruiters are concerned that candidates and hiring managers alike will be sceptical to the use of games as a medium for administering tests, and also to the recommendations made by the algorithms.

The focus group participants were positive to including such game-based solutions as part of their traditional processes yet limited to the hiring of specific categories of candidates (typically for junior positions). They did, however, not see an imminent threat in Pymetrics or Knack or similar providers disrupting the business of recruitment.

If my informants are representative of the recruitment industry in general, the risk of disruption is genuine. The recruiters do not consider the newcomers as threats in the high-end market for recruitment services, and I think they are right. The problem is that entrants don't need to attack a market from the top in order to disrupt it. Classical disruption occurs when new competitors offer cheaper services to clients at the low-end, and move upwards from there. Such is a likely scenario to play out in the case of gamified recruitment.

CONTENTS

CHAPTER 1: INTRODUCTION **1**

CHAPTER 2: LITERATURE REVIEW **2**

RECRUITMENT AND SELECTION 2

GAMIFICATION 7

GAMIFICATION OF RECRUITMENT AND SELECTION 8

CHAPTER 3: METHODOLOGY **14**

CHAPTER 4: RESULTS **18**

TWO CASE-STUDIES ON GAMIFIED PSYCHOMETRICS 18

FOCUS GROUP WITH INDUSTRY EXPERTS 30

CHAPTER 5: DISCUSSION **42**

INCREASED PULLING FORCE 42

BETTER SELECTION 42

GAME CHANGERS? 43

DISRUPTING FROM BELOW 44

GAMIFIED ASSESSMENT PLATFORMS 45

SYSTEMIC CHALLENGES TO GAME-BASED ASSESSMENT PLATFORMS 47

BASIC REQUIREMENTS 49

LIMITATIONS AND FURTHER RESEARCH 53

CHAPTER 6: CONCLUSION **54**

CITATIONS **55**

CHAPTER 1: INTRODUCTION

Apparently, businesses find themselves in a "war for talent", a fight that is supposed to have been raging in the Western world for more than two decades (e.g., Michaels et al., 2001)[i]. Others claim the war to be over - and that the *talents* won[ii].

War or not, talent acquisition is now said to be the third-most-important challenge companies face, with 81 percent of respondents in a recent report calling it important or very important (Deloitte, 2017)[iii].

In the two-tier job market that has emerged over the last decades (Searle, 2009), several groups of specialists find themselves in high demand and in the state of being continuously courted, making The Wall Street Journal refer to the deprived recruiters tasked with unearthing and convincing them to apply for positions as having "the loneliest job in a tight labour market."[iv]

Responding to that situation, the traditional recruiters work harder by the day to satisfy the needs of their best-paying clients. In today's job market, a unique selling proposition that many recruiters emphasise is how they provide their clients with access to candidates through their own personal networks. Another strength lies in their ability to assist the hiring managers in the subsequent procedure of selection, by their mastery of scientifically proven tools. Many of these tools are labour intensive, and so is the process of sourcing via the recruiters´ personal networks. The bill sent by recruitment agencies can typically reach NOK 300.000 (US$ 36.000) per hire (focus group, 2018).

At the same time, the $200 billion a year recruitment industry is starting to attract the serious attention of tech startups. They offer a new take on the processes of talent attraction and identification, with products and services ranging from digitizing the steps involved in traditional hiring, to completely changing these practices through the use of emerging technology.

The providers of game-based psychometric assessments for personnel attraction and selection are of the latter category. A common denominator among them is their ambition of reducing the value of *background*, academic or other, to the benefit of the candidates´ *potential* to perform successfully in the hiring organizations. If these companies succeed in bringing out their own potential, they could become game-changers for hiring businesses in general, and not least for recruitment agencies.

The present thesis will look into two solutions of game-based psychometric assessments for personnel attraction and selection, provided by the US-based companies Knack.it Corp and Pymetrics. I will hold their "Netflix like" approaches to hiring up against the existing practices within recruitment and selection, as described by HR literature and by a focus group of recruitment industry experts. I will relate the findings to the context of gamification in general. Moreover, I will look into the effect that the introduction of such tools could have on the market for recruitment services.

Based on this, the questions that I would like to answer through the present study are:

"Which are the potential benefits and limitations of gamified psychometrics for recruitment and selection? How likely is the prospect of game-based assessment providers disrupting the recruitment industry?"

CHAPTER 2: LITERATURE REVIEW

I'll first provide an overview of the field of recruitment and selection, followed by a brief description of gamification as described by scholars during the decade since the term emerged, and lastly look at some of the available literature combining the two fields; gamification of recruitment and selection. The purpose is to give a backdrop to a discussion about the function of the new game-based recruitment tools and the processes related to these versus the classical recruitment tools and processes.

RECRUITMENT AND SELECTION

What is meant by the terms of 'recruitment' and 'selection', and which are the key processes and tools related to these undertakings?

DEFINING RECRUITMENT AND SELECTION

The goal of the processes of recruitment and selection is to attract, identify and retain staff for an organization.

Hereunder, the process of *recruitment* is defined as the attraction of capable candidates to a vacancy.

Following recruitment comes *selection,* which concerns the assessment and identification of the suitability of the attracted candidates. (For all of the above definitions: Searle, 2009[v])

I´ll now go more deeply into the two processes of recruitment and selection, respectively.

Recruitment

Attracting enough relevant candidates for a position is obviously fundamental. As such, you can judge the overall utility of any recruitment and selection system by assessing the quality of the applicants attracted (Carlson *et al.,* 2002 in Lievens & Chapman, 2009)[vi].

This need to attract capable candidates makes the employer undertake various actions that in sum are intended to:

1. bring a job opening to the attention of potential job candidates;
2. influence whether these individuals apply for the job in mention;
3. affect whether they maintain their interest in the position until an offer is made, and not least;
4. influence whether a job offer is accepted (Barber, 1998)[vii]

Motivation is key: The tools used must ensure that the right candidates feel the drive to present themselves as candidates for the job, that they stay in for the whole process, and that they take the job offer when it's given. That is worth noticing when we´ll later look into the concept of gamification, which typically is described to be precisely about motivation and engagement (Deterding, 2012)[viii].

A company is only as good as the people working for it, and the organization´s recruitment processes is vital to bringing in the right candidates. This is hardly a new situation, but it's likely that the importance of recruitment has increased in the last decades. "A two tier job market has emerged over the last few years with increased competition amongst employers to attract the most qualified applicants (Michaels *et al.*, 2001 in Searle, 2009). As a result, attracting the most qualified candidates is more critical to organizational success (Chapman *et al.*, 2005 in Searle, 2009)."

How is this need to attract capable candidates met by recruiters and employers today? The literature describes job advertisements, whether on corporate web sites, in traditional media, social media or on job advertisement platforms, and word-of-mouth, the latter increasingly often generated at scale by posting employee testimonials on the internet, as the most common contemporary recruitment tools, in attrition to a relatively new method, the use of "bots" to search for candidates through public on-line information, a trend which is increasing (Searle, 2009).

An interesting fact to bear in mind when discussing recruitment is that very few people are active job seekers. It's well-known that the unemployment rate is relatively low in the Western world, and even on online job sites, only 5 % of visitors are unemployed. The great majority (71 %) of are in fact employed people who are not actively thinking about changing jobs. Only 10 % of the remainder are actively looking for a new job. (Bartram, 2000)[ix]

As a consequence, in order for a recruitment process to have a broad impact, the recruiter needs to convince already employed people who are not actively looking for a new job, to do exactly that. And the goal is not to convince merely any employed persons, but specifically the right ones for the open position. There's no shortage per se of candidates on job boards, the problem is finding the ones that might be relevant: "We have to sift through lots of résumés," as an executive interviewed by Forrester put it, "kissing frogs before you find the prince" (Lawrence, 1999 in Bartram, 2000). [x] At the turn of the last century, Bartram envisioned a future where "new technology is used to apply valid objective assessment techniques to the initial sift process." By doing this, he claimed, you could match the competencies and capabilities of the applicant to the requirements of the job vacancy, "and so produce a high quality shortlist that only contains princes." (Bartram, 2000). We´ll later see if gamified assessment tools have brought the vision of a high prince-to-frog ratio closer to reality.

Furthermore, it's common among managers to express the wish for candidates to not only be personally fit for the vacant role, but also to help increase the diversity in the organization´s workforce. In a survey of nearly 9,000 recruiters and hiring managers from 39 countries, 78 % responded that increased diversity was very or extremely important to them. The biggest barrier to improving diversity, according to their experience, was finding diverse candidates to interview. (LinkedIn, 2018)[xi]. Diversity strands on recruiting. Could the introduction of game-based recruitment lay the grounds for a more diverse workforce? That's a point I'll be returning to later.

Selection

After (hopefully) having attracted a number of candidates in the recruitment phase, the task of picking the right one remains. In the following I'll outline some of the most common tools used to that end today, in an attempt to present the context into which the game-based assessments are introduced. For the same reason, I'll also review what research says about the relationship between the respective tools and the subsequent job performance of the selected candidates, the so-called criterion related validity co-efficient.

Interviews

Interviews remain the most common method of selection (Searle, 2009). The reported relationship between interviews and subsequent job performance varies greatly, from r = 0.19 to r = 0.62 (Schmidt and Radar, 1999 and Wiesner and Cronshaw, 1988 in Searle, 2009). Interviews are typically defined as either structured or unstructured. The use of structured interview formats returns the highest average validity, with r = 0.51 (Robertson and Smith, 2001 in Searle, 2009). In structured interviews, recruiters are asking "pre-defined questions and assessing responses against pre-defined response categories" (Huffcutt and Roth, 1998 in Searle, 2009).

If you search the web for "job interview is dead", you're likely to be over-whelmed by problems related to interview as a selection tool. Eg.: "(…) we are incapable of treating the interview data as little more than unreliable gossip.»[xii], "I say, throw out the interviews altogether»[xiii], and "Will interviews go the way of the dinosaur?"[xiv] There is also mention of other tools and techniques that could replace the job interview, amongst those, game-based assessments. In that perspective it could be worthwhile noticing how some scholars emphasise that the function of the interview goes beyond pure selection. There is an important social aspect related to the interview, where both parties get to meet and formally assess each other (Searle, 2009).

Cognitive ability tests

Cognitive ability testing is arguably one of the most effective means of selecting to enhance job performance (Searle, 2009) which is probably why cognitive ability tests appear to be the cornerstone of most game-based candidate assessments I've looked into.

Cognitive ability tests are designed to assess performance, ideally the maximum performance the candidate can currently achieve, typically by providing multiple choice items based on numerical and comprehension skills (Searle, 2009). The job performance for validity for cognitive ability tests alone is around r = 0.51 (Robertson and Smith, 201 in Searle, 2009).

The literature makes mention of some potential pitfalls in the use of cognitive ability tests. Such tools might purport to measure so-called fluid intelligence, which is the application of reasoning skills into novel problem-solving situations, while they in reality are mixing in the measurement of so-called crystallized intelligence, which are reasoning skills more linked to the candidates' culture and/or education. If that's the case, it could leave the candidates' true ability un-assessed. It's a concern for selection tools that rely greatly on cognitive ability tests that these differences could be systematic. Individuals

from ethnic minority groups have been found to perform below average on such tests (Bobki et al., 1999; Scmitt and Mills, 2001 in Searle, 2009), possibly for the reason that they measure crystallized intelligence. Older people have also been found to use culturally-specific skills in problem-solving more often than younger people (Horn and Noll, 1994 in Searle).

However, for the possibly motivation-boosting game-based assessments, it's noteworthy that differences in motivation could also reduce the validity of cognitive test results tests, possibly enough to explain the race-related difference in the results (Ployhart and Ehrhart, 2002 in Searle, 2009).

Another issue is that tests are frequently designed without validating them to on-the job performance, but to other general intelligence instruments, possibly due to "statistical convenience" (Kline 1993 in Searle, 2009), a topic I'll return to in the discussion.

Personality tests

Personality tests are used to discern candidates' behavioural style and disposition (Searle, 2009). The reported predictive validity ranges from r = 0.21 to r = 0.40 (Schmidt et al., 1984; Robertson and Smith, 2001, in Searle). The use of personality tests for selection was limited until the 1990s, when new studies described their utility. After that, the scholarly interest in personality research related to selection or job performance increased dramatically. (Morgeson *et al.,* 2007). [xv]

The two game-based assessments I've reviewed for this study include elements of personality assessments. An example is their adoption of the Balloon Analogue Risk Task (BART), a computerized measure of risk-taking behavior (White *et. al.,* 2008)[xvi].

Work sample tests

Work sample or situational judgement tests (SJTs) are used to assess the candidates' judgements, abilities and behaviours in the context of a specific work situation (Searle). The job performance validity is reported to be consistently high, with r = 0.56, increasing to r = 0.60 when combined with a cognitive ability test (Robertson and Smith, 2001, in Searle). Searle suggests that situational judgement tests have the advantage of testing for cultural fit (*how* task are performed) and not only skills and abilities (*what* is done).

For their increased validity, and the possibility of scaled delivery of such tests via the internet, the use of SJTs is increasing rapidly, "replacing multiple-assessor, multiple-test approaches, or are being used in pre-screening" (Searle, 2009). A downside to SJT is the extra cost of producing them, as they need to be tailor-made to each work context. Also, results from SJTs have been found to have lower durability that for more generic ability tools (Siegel and Bergman, 1975 in Searle, 2009).

One of the strengths of computer games is their ability to bring the players into a new environment in an immersive way: "A main reason to play computer games is the pleasure of being immersed in a mediated world.»[xvii] For that reason, gamified recruitment and selection tools could be apt as media for delivering SJT, a proposition that'll be brought up in the discussion chapter.

Integrity tests
Integrity tests are intended to predict undesirable work behaviours such as theft and absence. Their predictive value is high, in particular when coupled with ability tests, with R = 0.65 (Robertson and Smith, 2001, in Searle, 2009).

When I mention integrity tests as a selection tool, it's not because of their relevance to the gamified psychometrics that I've studied, although it could be. The reason is to typify the gap between theory and practice amongst recruiters and recruiting managers, since integrity tests represent a method that is used to a small extent, despite the reported high validity. "Senior HR professionals appeal to either ignore or remain unaware of clear scientific findings" (Rynes et al., 2007, in Searle, 2009), resulting in the limited adoption of the selection tools with the highest job performance validity. Recruiters and hiring managers often choose their tools based "on ease and familiarity – and rarely correlate them to results." (Bersin and Chamorro-Premuzic, 2019)[xviii]

An extreme example of selection tools chosen for familiarity and not for proven validity can be found in France: 75 % of French employers have been found to use *graphology* as part of their recruitment process[xix] even though the validity of the method is reported to be only .02[xx]

This is significant to keep in mind when judging the real-life potential of new, scientifically based tools for recruitment and selection. Another aspect to consider when judging the global market for such solutions is that, "evidence for the universality of recruitment selection exercises is limited" (Lievens and Chapman, forthcoming in Searle, 2009). That could constitute a challenge for tools who rely on scalability.

GAMIFICATION

Defining gamification

Gamification is a relatively new word. According to some sources, the first documented use dates back to 2008, but the term did not see widespread adoption before the second half of 2010. (Deterding *et al.,* 2011)[xxi]

The term is mainly used to describe two concepts, distinct but still related. The first use concerns the fact that games have taken an increasingly large part in everyday life for an increasingly large part of the public (Deterding *et al.,* 2011). According to Nielsen, "In 2018, two-thirds of the U.S. population 13 years and older are (self-identified) gamers, up from 58% in 2013.»(Nielsen, 2018)[xxii] This trend of the increased impact of games could be referred to as *gamification of society*: "This century has been called the 'ludic century', as games and play have become the dominant cultural form of the era."[xxiii]

The second concept to which the term *gamification* is applied deals with the use of game elements in attempts to make non-game products and services more enjoyable and engaging. Following on that, gamification can be defined as as the use of "elements from game design in non-game contexts, products, and services to motivate desired behaviours." (Deterding, 2012) [xxiv]

A 2014 review of 24 empirical studies on gamification[xxv] mentions the following different contexts where gamification is used:

- Commerce;
- Education/learning;
- Health/exercise;
- Intra-organizational systems;
- Sharing;
- Sustainable consumption;
- Work;
- Innovation/ideation;
- Data gathering.

The motivation for using game design elements in these and other non-game contexts would vary. But likely, the providers of gamified solutions and the managers buying them have thought like Sebastian Deterding: "Games entice hundreds of millions of people across the globe to spend countless hours and dollars performing often menial tasks – certainly, there must be some way to use this power for other purposes?" (Deterding, 2012)

The notion is that games make people do things they wouldn't have done, without the presence of the games. The idea behind gamification is to use of elements from game-design to generate new type of responses on the side of the users: "Gamification can be used as a tool to improve the participation and motivation of people in carrying out diverse tasks and activities that generally could not be too attractive" (Aparicio et al, 2012).

Does gamification work? Judging from its increased utilization, an increasingly large number of managers cross their fingers that it does: "The use of gamification has exploded in recent years." (Lowman, 2016)[xxvi]. The mentioned meta-study (Hamari *et al.,* 2014) describes that "According to a majority of the reviewed studies, gamification does produce positive effects and benefits." The study assessed the effect of gamification in the different above-mentioned contexts (such as for education/learning). In the following, I'll review some of the literature looking specifically at gamification in the context of recruitment and selection.

GAMIFICATION OF RECRUITMENT AND SELECTION

During the last few years, several ports have identified gamification as a top trend in the field of human resource management (e.g., Munson, 2013; Society for Human Resource Management, 2014, in Armstrong *et al.,* 2016[xxvii]). Moving on to the more specific, on the HR domains of recruitment and selection, I've highlighted some quotes from literature to illustrate the recent impact of this field:

> Technology presents unique opportunities to extend the boundaries of recruitment and selection processes, to truly open access to new applicants wherever they are. (Searle 2009)
>
> Gamification is an effective tool for winning the war for talent by providing HR practitioners an innovative tool to identify, attract and retain talent. (Lowman, 2016)
>
> Use of gamification in recruitment, onboarding, training, and performance management are on the rise in organizations as businesses turn toward technology to meet their objectives. (Callan et al, 2014)
>
> The increased use of technology in the deployment of ability and other selection test far earlier in the selection process has given recruiters and test developers access to a much wider range of candidate data. These new data have provided novel opportunities for the 'forensic' analysis of data streams. (Bartram, 2008)

In practical terms, how has the tool of gamification been set to use, and how has it worked, earning it this rapidly gained fame?

First of all, it's hard to claim that gamification is effective per se, "but specific game design elements have specific psychological effects" (Sailer *et al.,* 2016). It is not within the scope of the present thesis to assess specific game design elements used for recruitment and selection. In the following, I'll look at the available literature covering gamification for hiring, not covering possible differences in outcome stemming from difference in specific game design.

EXAMPLES OF GAMIFIED RECRUITMENT AND SELECTION

An early example of utilizing gamification to identify and attract talent is the free state-of-the-art PC game *America's Army: Special Forces,* released in 2002. "The game is developed and published by the US Army to attract new recruits" (Nieborg, 2004)[xxviii]. The game is broadly considered a success, "with numerous revisions and updates over the past 14 years." (Lowman 2016). One year after its release, the game, designed as "a realistic online multiplayer tactical squad-based First Person Shooter" (Nieborg, 2004), had more than three million registered accounts (Aarseth, 2003 in Nieborg, 2004)

Originally, *America's Army* was "first and foremost meant as an advergame, which refers to "the integration of advertising messages in online games." (Nieborg, 2004). Another frequently cited example on the use of gamification motivated by a wish of improving the potential applicants' attitude towards the recruiting organization, is the game *My Marriot Hotel*, by the hotel chain Mariott International. *My Marriot Hotel* simulates the duties of hotel managers. "By doing this they hope to change the attitude towards the hotel industry and Marriot in particular." Lopez, 2011 in Chow & Chapman, 2013[xxix]

Returning to *America's Army* and the initial intended purpose of the game: "Opposite to some advergames, the Army does so without gathering explicit information about gamers." (Nieborg, 2004). Nevertheless, by 2016 it's clear that "Individuals who advance quickly through the ranks in virtual training and skill development are likely to excel in real-life training and skill development as well. The Army can then target these high profile candidates for recruitment." (Lowman 2016). *America's Army* had moved from a pure advergame to be a tool for recruitment and selection.

Another early example of gamification used to identify talent is the game studio Electronic Arts' project "EA University". EA sought to identify top talent among its staff "by monitoring the speed at which employees complete challenging games and successfully collaborate with other employees participating in the games" (Lowman, 2016).

A more recent case of talent identification per gamification is *Wasabi Waiter* by tech start-up Knack.it Corp. Taking on the role of a waiter in a busy sushi restaurant, "The game evaluates every millisecond of player behaviour, gauges conscientiousness, emotional cognition, and other qualities which are said to be related with job performance. Apart from that, the app captures and analyses huge volumes of data developed by designers and behavioural scientists. The results pinpoint those strengths and skills which are ideal for hiring such as empathy or critical thinking. (…) The game then scores each player and calculates outstanding employees." (Joy, 2017)

Wasabi Waiter, which I'll be returning to as part of my case study on Knack, forms part of a "new breed of psychometric tests for recruitment" where "user´s choices and behaviour are mined by computer-generated algorithms to identify suitability for a given role" but doing so in a way that enhances the candidate experience above that of classical psychometric tests (Bersin and Chamorri-Premuzic, 2019). The latter article mentions three specific cases of such novel psychometric testing:

- HireVue´s MindX, used to predict candidates´ IQ through "sleek games";
- Arctic Shores, using a series of games resembling "1990s arcade games" to assess personality traits and competences; and

- Pymetrics. In a game that puts candidates in control over self-inflating balloons, employers can test candidates' impulsivity and risk-taking. I'll return to both this company and their balloons game in the case study on Pymetrics.

INTENDED EFFECTS BY GAMIFYING RECRUITMENT AND SELECTION

Why have hiring organizations started to introduce game-elements for recruitment and selection? The literature proposes some different motivations on the side of the employers:

"We propose that gamification in the recruiting context is essentially an attempt to change a potential applicants' attitude towards the organization through engaging them more deeply in the recruitment process (Chow and Chapman, 2013).

"Gamification opens the door for collecting information on an individual´s talent potential while also fostering organizational attraction and employee retention." (Lowman, 2016)

Notably, **enhancing the attractiveness** of the hiring organization to potential job candidates, and/or engaging new audiences, seems to be a recurrent motivation for the employers: "The nature of games (they are fun) enables their fluidity and propels them across social media outlets such as Facebook, LinkedIn, Twitter and etc. The potential for gamified recruitment applications to reach a wide audience is highly instrumental for organizations. This may be the defining factor of a gamified recruitment process" (Chow, Chapman, 2013).

I do agree with Chow and Chapman: Attraction is an essential part of gamification in the context of hiring. But whom you want to attract depends on whom you want to select, and gamified psychometrics could do both in one flow: "Candidate elimination is done at a much faster pace as it allows firms to test specific skills like time management, creative and innovative thinking etc." (Joy, 2017)[xxx]

As to **selection**, how do game-based assessments fare against non-game-based tests?

I have not been able to identify literature comparing the predictive force of gamified psychometric tests for job performance versus more classical psychometric tests. However, a study[xxxi] (Lappalainen, & Fanni, 2017) where "job candidates" in a mock selection situation completed game-based assessments for cognitive ability, in the form of "casual games", and traditional assessments, found that game-based assessments of cognitive ability have a medium to large positive correlation with traditional assessments of cognitive ability (r = .35 - .50).

That result must be seen in connection with lower costs per candidate tested, which is one of the obvious benefits from gamification used for recruitment (Wozniak, 2015). The relative validity of game-based assessments compared to traditional assessments of cognitive ability must also be linked to another recurrent motif in the literature on gamification of recruitment and selection: the increased motivation experienced by candidates taking gamified psychometric tests versus traditional ones. Lappalainen (2017) found that game-based assessments were associated with higher test-taking motivation (TTM) than the traditional assessments, and with lower anxiety than some of the traditional assessments.

If it's true that game-based assessments correlate with traditional tests, and they are cheaper to deploy at scale, they increase the attractiveness of the hiring organization, as well as the motivation for the test-takers, which other benefits could be present?

A frequently noted motivation for gamifying hiring processes, and one that the providers of such tools tend to emphasise, is the prospect of reducing *biases* towards candidates. There are several aspects to consider in that respect: One could be a possible increased fairness in gamified psychometric tests over traditional assessments. A study (Lappalainen, 2017) examined whether using games might reduce racial differences in performance because of lower perceived stereotype threat, higher test-taking motivation, and lower anxiety, but was not able to demonstrate such an effect, calling for further research on the subject.

As to gender discrimination, a white-paper from Pymetrics[xxxii] states:

> Women fare worse than men on standardized tests, multiple-choice type tests and self assessments. This gap does not stem from inferior competence, rather it stems from proven gender differences in test taking." For instance, "Research shows that on multiple-choice tests men tend to guess while women prefer to skip questions.

Pymetrics claims that with their solution,

> The pymetrics science-based game design removes the gender bias propagated by standardized tests and self assessments." Pymetrics describes[xxxiii] a case where, thanks to this effect and others, an un-named multinational financial services corporation was able to source in more gender neutral candidates: "The pymetrics recommendation pool consisted of 43% women as opposed to their traditional turnout of 31% women. The use of pymetrics increased the number of females in first round interviews by 19%.

A group of researchers at Pymetrics claim to know why "Hiring processes are still biased. (…) This is because recruiters, men and women alike, fall prey to unconscious bias." (Narayanan *et al.,* 2016)[xxxiv] They describe the state in the top five orchestras in the U.S in 1970, with fewer than five percent women employed, and the novel solution that was launched to change that:

> In the 1970s, a number of orchestras adopted "blind" auditions whereby screens are used to conceal the identity and gender of the musician from the jury. In the years after these changes were instituted, the number of women increased to 25 percent, and two decades later, the number was at 35 percent. Gamified assessments can serve as a form of blind auditions.

To me, it is clear that gamified assessments can serve the function of blind audtions, but I've not found independent studies intending to demonstrate such an effect.

ISSUES OF CONCERN AND CRITICAL VOICES

Although referring to recruitment via the internet, before the concept of gamification was coined, Bartram (2000) lists some issues of good assessment practice that could be relevant also for game-based processes:

- Security and confidentiality: By administering computer-based test to large groups you end up with large amount of highly sensitive data. How to ensure the security of this data, and to prevent unauthorised access to it?
- Authentication: How to ensure that the test taker is the real candidate, taking the test unaided?
- Controlled test conditions and equal access: Do all test takers have sufficient quality of Internet service and equipment needed for the test?
- Control over practice: Do all test takers have a similar amount of practice for the test form?

I've had this check-list in mind when reviewing the two cases, Knack and Pymetrics, and will return to some of these points later in the discussion.

Game-based candidate assessments are generally delivered over the internet, and not in supervised settings. That gives a cost-effective high-volume administration. But as the use of remotely delivered tests is increasing, so is the concern about faking and cheating (Searle, 2009). Several techniques are now implemented to reduce the damage of fraud while maintaining the benefits of large scale, such as photo verification systems to ensure that the real candidate is the one performing the tests (Searle, 2006 in Searle, 2009).

Constituting another critical voice on the use of game-elements in the selection process, Woźniak (2015) states that

> Playing a game is not equivalent to taking a part in an Assessment Center – even if events or tasks in the game reflect those that are critical for the job; the actor's behavior in the game is not necessarily indicative of how s/he would behave in everyday life (Woźniak, 2015). Playing the game has its own dynamics, and how players carry out their roles is only a (weak) approximation of how they would perform their professional roles.

I believe that Woźniak is here relating mainly to game-simulations of work situations (gamified SJTs), and not, for instance, gamified cognitive ability tests. If so, that concern is less relevant to the two cases I've studied.

As games for recruitment and selection are administered over wide geographical areas, local particularities need to be recognized and considered. (Shuler *et al.*, 1993 in Searle, 2009). These could be differences in acceptance of faking and cheating (Searle, 2003 in Searle, 2009) and other differences, for instance in the emphasis that employers place on proven job experience (strongly in Australia, Germany and the US) versus innate potential (strongly in South Korea, Taiwan and Japan) (Huo *et al.,* 2002; Von Glinow *et al.,* 2002 in Searle, 2009). In general, evidence for the universality of recruitment selection exercises is limited (Lievens and Chapman, forthcoming, in Searle, 2009).

Searle further raises the concern that the new tools might not assess the skills of certain groups such as older workers with a similar validity as for younger workers. Although not focusing specifically on recruitment, a study on demographic differences in perceived benefit from gamification found that "ease of use of gamification is shown to decline with age». Searle warns that "we need to ensure that we are identifying real job performance differences, not creating new barriers to entry." (Searle, 2009)

CHAPTER 3: METHODOLOGY

The book Prediction Machines: The Simple Economics of Artificial Intelligence (Agrawal *et al.,* 2018)[xxxv] gave me the direction for the present thesis. Artificial Intelligence (AI) is all about prediction, and the process of recruitment and selection stood out as the ultimate prediction task. At the same time, the recruitment industry appeared me to as an old-fashioned business, with labour intensive processes and a notably low degree of automatization. My initial idea was "AI recruitment", and the first searches on that subject yielded some interesting examples of start-ups that seemed eager to help transform the recruitment industry. I found at least 25 [xxxvi] different products and services using AI is some way for the process of assessing candidates.

I later discussed the subject of AI recruitment with Paul Daugherty, CTO at Accenture (a consultant firm), after a guest lecture he held in a May 2018 class in the curse "The global business of AI and robotics" at MIT Sloan School of Management. He mentioned Knack and Pymetrics as two start-ups finding themselves in the forefront of AI recruitment – and specifically in their use of *gamification*. Using game-based elements to attract and select candidates arouse my interest, possibly due to my prior experience working with computer game developers. The emphasis placed by these two providers on reducing the effect of human *biases* on the outcome of recruitment processes also struck me as exciting. I wanted to learn how these services might affect the work of recruiters, how they could affect the recruitment industry as a whole, and how the industry would react to that. Could we be talking of true *game-changer*?

As mentioned, Pymetrics and Knack were not the only players in the market for candidate assessments with some element of artificial intelligence. Neither were they the only providers of game-based assessments. An article[xxxvii] mentions the companies MINDX (later merged with HireVue, Scoutible, Artic Shores, Cognisess and Cut-e in addition to Knack and Pymetrics in the context of game-based candidate assessments, as per the third column here:

I found that several of the above-mentioned solutions could've served well as cases for the present study. When I decided to go for Pymetrics and Knack, it was because they appeared to be relatively well-established players in the market, with similar but also somewhat different approaches to gamification. Last, I found them open to providing access to their services and sharing their insight. I further decided to limit the selection

to precisely two companies, in order to keep the scope of the study manageable. Or else, Arctic Shores would probably have been my third choice.

An alternative explanation for why I stuck with Knack and Pymetrics, could be a well-known phenomenon from precisely recruitment and selection, namely *confirmation bias*. As these two had been mentioned by an authority in the field prior to my investigation, my impression was already made up, and that could've influenced me enough to rule out the possible competition. Maybe an algorithm could have done a better, in the sense of a less-biased, job in that sense?

Having selected the two research subjects, I needed to find out how they worked and not least how to gather data on them. A quantitative study, comparing the results from a game-based recruitment and selection process versus a traditional one, was the first option I reviewed. I Soon I concluded such a study to be outside the practical scope for this thesis (six months of writing, in parallel with a full-time job).

As to the option of basing the thesis on secondary data, both Pymetrics and Knack had published case-studies on different utilizations of their products. But they were all written by the companies themselves (or with a strong contribution from them), and as such not independent studies suitable as sufficient sources to describe the two cases.

For the above-mentioned reasons, I decided to gather primary data on the two services by accessing them myself, taking on the role of a future job candidate and also that of a recruiter (or an HR manager at a recruiting organization.)

Gaining access to the two services proved to be relatively straight-forward. For Knack I had to submit a research proposal form, which was approved 2,5 weeks later, on the condition that I would "keep us (the company) in the loop with your research, as well as the data that you collect (we use these research projects to refine our app"). Accepting these terms, I received a link to Knack´s "Business app", allowing me to set up a fictitious recruitment campaign. Sending the link to myself, I was then able to test the solution from the candidate perspective. Subsequently I could access the data in that campaign, like if I were on the recruiter's end. Due to my acceptance of Knack´s research terms, I could access the campaign data for free, whereas a "normal" business would´ve had to pay to see the results (paying a fee per candidate).

In the case of Pymetrics, I got in touch with the company by stating my interest through a form on their website. That initiative generated an email from a sales manager, and the next day, I received a link to register as a candidate. A couple of weeks later, I followed up asking the same Pymetrics representative if it would be possible to access the employer/recruiter side of the solution, and not just the candidate´s. That query was left unanswered. When I chose not to submit reminders, it's mainly because I in the meantime had learned that the candidates in Pymetrics' solution themselves get access to a comprehensive and quantified "trait report", whereas you'd need to be on the recruiting side of Knack to see something similar. Another reason for not insisting on accessing Pymetrics' business platform, is the manual process needed to set up a campaign in their solution. While Knack could be pretty much "plug-and-play" for a new business client of theirs, Pymetrics' starting point is usually a data analysis of the recruiting organization's needs. That would obviously have been harder to simulate for the purpose of this study. I thus concluded that the access given to me be would satisfy my need for sufficient insight into both services.

Playing through the games was the next step, first Knack's, then Pymetrics', documenting to the best of my abilities along the way. For Knack, I used an iPhone (X) to download and use the KnackApp. Pymetrics was tested using an iMac. I took screenshots of the games during progress (mostly stills in the beginning, mostly videos in the end, since it proved difficult to simultaneously play and take photos), and I made notes describing the games in the brakes between them. I tried to describe not merely the content of the games but also their context: What did it look like the game developers were aiming to assess in each game? And which other tests, that I had seen as a manager or as a job candidate myself, did the games resemble? Also, what was my user experience like?

As to the recruiter or business side of the systems, I recorded the scores that Knack gave me from my own test. I also noted the results from their "Demo campaign": All Knack´s business users are free to share the link from this demo, and as of the present, 54 users have played the two games in that campaign. I took screenshots from their results, anonymizing names and their profile pictures as well as their email addresses to protect their privacy.

Having documented the two services, I was keen to present the findings to experienced recruiters, in order to get their reactions and thoughts on the potential implications these new tools could have on their industry. A fellow student at the NTNU Master in Technology Management class of 2019 gave me a suggestion for a candidate for such an interview: Senior Executive Researcher at Mercuri Urval Norway´s Bergen office, Tarald Eik Mong. I met him and presented the outline for the thesis, without sharing any details on the two cases I had studied. After the meeting, I thought that assembling a focus group could provide useful further insight into the perspective of the industry.

Following Robert K. Yins definition of a focus group, "A form of data collection whereby the researcher convenes a small group of people having similar attributes, experiences, or "focus"", (Yin, 2016)[xxxviii] the idea was to assemble a group of industry experts that would have similar but complementing views on the concept of gamification of recruitment and selection. The group members needed to be able to reflect openly on the implications of such a novelty. Ideally, they should not fall into the category of *innovators* nor that of *laggards*, to use Everett Rogers terms (Rogers, 2003)[xxxix] related to people´s willingness to adopt ned technology. Rather they ought to represent the *early adopters or early majority* within their industry. That would provide a more relevant and realistic insight into the true business potential for these new tools.

Tarald seemed like a great fit for such a focused group. He agreed to participate and pointed me in the direction of two other possible candidates for that group: Hege Rødland, partner and manager at the Bergen office of Visindi, and Andreas de Lange, Recruitment Manager for IT in Bergen & Stavanger for Experis Norway. They both agreed to join, but sadly, Hege had to cancel the same day as the first (and only) focus group meeting. In parallel, I had met Hilde Lekven, Group CEO at OTIGA Group AS and Personalhuset Staffing Group. She expressed her interest in the subject and willingness to join the focus group, but we were not able to find a date that would accommodate her schedule. On Tarald´s advice, Anders Kleppe Norheim, Executive Researcher at Mercuri Urval Norway, was brought in to the focus group, meeting in December 2018.

Prior to the focus group meeting, I had prepared a PowerPoint presentation of Knack and Pymetrics (attached hereto), and a brief interview guide (also attached) for the subjects I wanted the group to discuss, as well as the questions that I needed answered. Apart from the that, the plan was to lead the group "in a nondirective manner" (Yin, 2016).

Over the course of 2,5 hours, I first showcased the two solutions, Knack and Pymetrics, as per the presentation, and answered specific questions from the group. I then set up the discussion, as per the interview guide, recording the conversation on a mobile phone, interfering or otherwise influencing the discussion as little as possible. After the meeting, I transcribed the conversation, translating it simultaneously from Norwegian to English (without me being, I should note, a certified English translator). The audio file containing the recording from the meeting was then deleted.

Assessing my own "research lens" in relation to this study (Yin, 2016), I tried to be aware that my own knowledge, views and predilections might influence the data collection. I realize that my initial view on the recruitment industry as old-fashioned and low-tech might have made me more positive towards high-tech solutions challenging the status quo of recruiters. My interest in computers games not only influenced my choice of research question but potentially also my view on the solutions studied, hypothetically making me both more positive towards them and more critical.

Furthermore, I'd like to point out a possible weakness in the way the focus group was formed. Those that (for various reasons) ended up in the final group were all male and in their thirties. Adding to their uniform profile, they had all studied at the University of Bergen (Norway) and they all worked or had previously worked at Mercuri Urval. I would obviously have preferred a more diverse group, but practicalities came in my way. I should add that the focus group members did not express uniform views, with the notable exception of one aspect, their lack of openness to the disruptive potential of services such as Pymetrics and Knack. That is likely more related to the group they were selected from than to their own personal characteristics:

All three focus group members work in the "high-end" of the recruitment market, filling positions in cases where the main problem is getting enough relevant candidates, and businesses outsource the recruitment to them in order to unearth potential applicants that the businesses would've had greater problems finding. As such, the group members are not presently occupied with the more "low-end" part of the industry, where the main problem is administering and selecting from a long list of applicants. Norwegian companies at the moment tend not to outsource the selection of applicants for such positions. The members of the focus group find that the willingness to pay for recruitment services is inversely proportional to the number of available candidates for any position, and that tints their view on the business potential of the game-based services. They rightfully don't see them as direct short-time threats to their own work. Their shared position within the recruitment industry might however make them blind to another possible angle of attack from the newcomers – from below.

That's a point I'll be returning to in the discussion chapter.

CHAPTER 4: RESULTS

The findings are related to two sets of sources, the former of these being case-studies on the two providers of gamified psychometrics, Knack.it Corp and Pymetrics, and the latter a focus group with representatives from the recruitment industry.

TWO CASE-STUDIES ON GAMIFIED PSYCHOMETRICS

KNACK.IT

Silicon Valley based Knack.it Corp. was incorporated in 2010[xl] by Israeli entrepreneur and Harvard doctorate Guy Halfteck. Apparently, he got the idea for the startup after getting turned down for an early-career leadership program at a top New York hedge fund. "That experience led Guy to look for a better way to showcase one's talent than the traditional credentials, pedigree, and interviews."[xli]

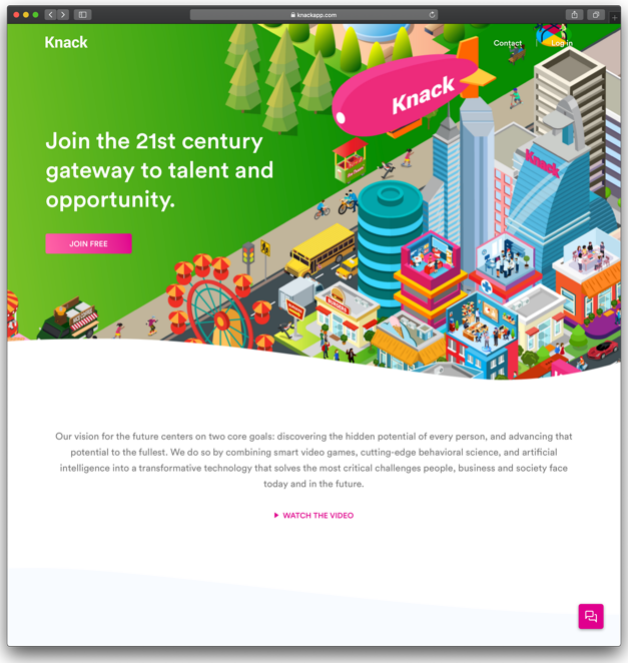
Knack refers to itself as a both an assessment tool and a platform matching candidates and employers "*by combining smart video games, cutting-edge behavioral science, and artificial intelligence into a transformative technology that solves the most critical challenges people, business and society face today and in the future.".* [xlii] [xliii]

The basis for the candidate assessment and the matchmaking with employers is a analysis of the recruiting organization's true needs:

"Knack will first help companies learn what drives success, and then use that to identify the right people for positions. (…) It's not just about intelligence, but rather the sum total of the human condition."[xliv]

For the purpose of this thesis, I've not tested the side of Knack's services occupied with "first help companies learn what drives success". Nor is that a mandatory way into the platform, as of the present; In fact, the marketing seems first and foremost directed towards the assessment of candidates (internal or external) early in the process.

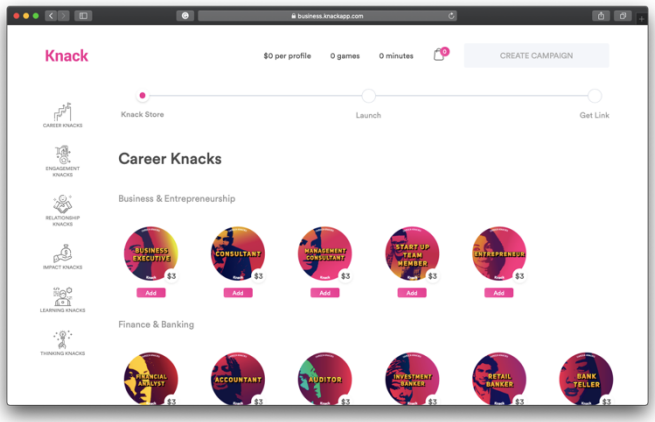
Knack's landing page is oriented towards two groups of users: "Individuals" on one side, "Employer", "High School", "Skilling Provider", "University", "NGO" and "Government" on the other.



As an individual, you're free to directly download the KnackApp (via App Store or Google Play), without prior registration on the web site. "The games are free to download and play. And it's free to see your results and share them.»[xlv] The app is available for mobile devices only.

As any employer, you could be just a brief registration and a few clicks away from setting up your own recruitment campaign, ready to be launched from Knack's platform.

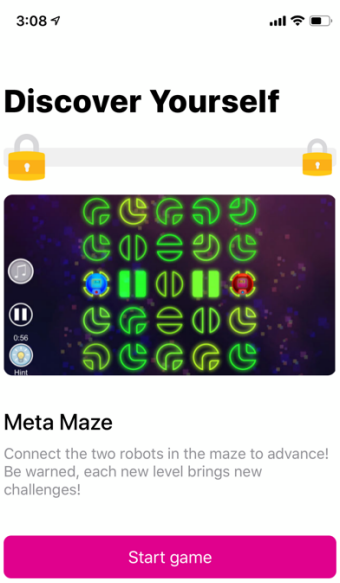
After registering a company profile, the person recruiting starts by choosing from a list of "knacks" that seem relevant for the open position(s). The "knacks" are defined traits, abilities and talent such as "learning from failure", "solving problems" or "taking someone else's perspective". There is also another list to choose from, the so-called "career knacks" or "super knacks". To me they appear to represent the same traits and abilities as the "individual knacks" , but bundled together and labelled to match the typical traits and abilities that a candidate need to have a high potential for a given career, such as, "financial analyst", "engineer", "sales associate", etc.



As an employer (on the *pay-as-you-go* plan), it's free to set up a campaign. After choosing the traits and abilities for which you'd like to test, a link to your campaign is generated. This link can then be shared to candidates through the medium of the hiring organization's choice (for instance Facebook). Candidates clicking on the shared link will be forwarded to their store for mobile apps, where the game can be downloaded. As mentioned, anyone can download the app without a connection to a recruitment

campaign. But unlike individuals downloading the game at their own initiative, using the link to access the game will make its content reflect the "knacks" that the employer wants tested.

I first tested the app using the link generated by Knack's pre-defined "Demo campaign". I also set up other campaigns, for a variety of different imagined positions. Summing up, I find that there are basically three games available on the platform. Depending on the specifics of the campaign, these could be either one, two or all three of the following games: "Meta Maze", "Balloon Brigade", and "Dashi Dash". In the following. I'll briefly describe my experience with the three games, and my understanding of the purpose of the different gamified tests, from my layman´s point of view (I should stress that I don't have a background in psychology).



The player's task in "**Meta Maze"** is to connect two robots by rotating different objects placed in the space between them, forming an open path between the two. You have 60 seconds to finish each level, after which you receive a score (1-3 stars) and pass on to the next level. There is a total of ten levels. The complexity rises rapidly, by increasing the number of objects that require modification in order to connect the start and end points, with the introduction of interdependently rotating objects, hidden objects and more.

The game seems to bear resemblance to some classical psychological tests, such as the Porteus Maze test: *"Mazes in general are thought to assess selection, trial, rejection or adoption of alternative sequences of conduct or thought. Porteus asserted that, like the Binet-Simon scale, it is a valuable supplement in evaluating subjects' foresight and planning abilities."*[xlvi]

I do see that "Meta Maze" could be apt for testing foresight and planning abilities, as well as memory, ability and willingness to learn from mistakes, and possible more. These being important traits for workers in quite a few different positions (judging from my own experience as a manager), it's no wonder that Knack brought up this game more frequently than any of the others. It appeared in almost all of my campaigns, for positions as diverse as: "Driver", "Barista", "Lawyer", "Architect", "HR Manager" and "Engineer".

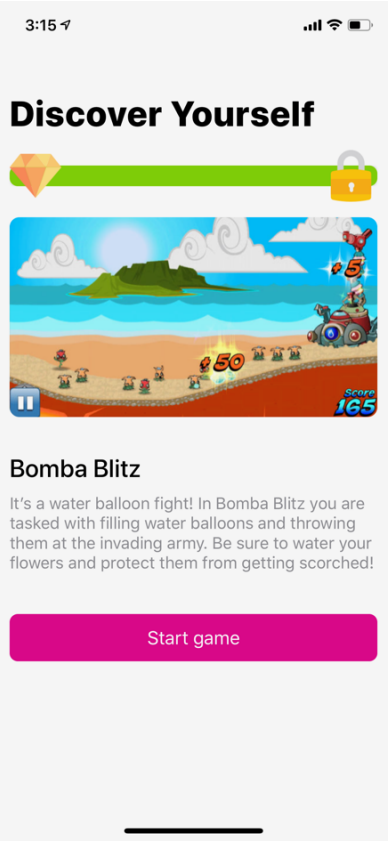
The second most common game generated by my selection of skills and professions is "**Dashi Dash"** (in an earlier version known as "Wasabi Waiter"). Here, the player is given the role of a waiter in a sushi restaurant. It rapidly gets more crowded, and you're expected to please the rather demanding clientele. The customers' facial expressions convey certain emotions, such as happiness, anger, and disgust, and need to be served a sushi dish labelled accordingly. You start by taking their orders, then prepare their dishes, which should be served to the right client before they leave angrily (the clock is ticking). You also need to bring the used dishes to the sink, or else, the flow of new customers to the restaurant will try out. With each level of the game, new emotions to be read and acted upon are introduced.

I would say that "Dashi Dash", like "Meta Maze", is about testing the player for foresight and planning abilities, as well as the ability to cope with stress and pressure.

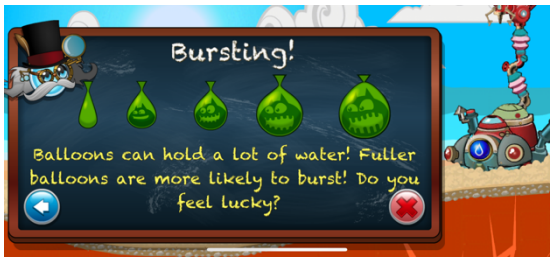
The uniqueness of the sushi game specifically lies in its testing for emotional cognition based on the player's success in recognizing certain emotions in (drawn) faces. Measuring someone's ability to read nonverbal social information is not a novelty by itself. There are several tests in the market for measuring the perception and identification of emotions from faces, with Brief Affect Recognition Test (BART) and the Japanese and Caucasian Brief Affective Recognition Test (JACBART) as the most frequently used.[xlvii] The way that Knack has designed "Dashi Dash", as a simulation game is the most apparent differentiator.

In the campaigns that I set up, "Dashi Dash" showed up in tandem with "Meta Maze" for the positions of "Barista", "Lawyer", "Architect", "HR Manager", and solo in the cases of "Account Executive" and "Business Executive".

The least common game to show up as a result of my selection of "super knacks" was "Bomba Blitz". Here you are the commander of a weapon capable of filling balloons with water and propel them through the air, towards the approaching invaders. Unless challenged in time, these attackers will set on fire the flowers you are set to defend. In idle moments, of which there are few, you can gain extra points by watering the mentioned plants. However, the most important decisions for the player are related to the amount of water to put into each balloon before sending them off against the invador. Too much water, and it'll burst. Too little, and the impact of the water bomb will decrease.



"Bomba Blitz" is not the first example of a test using balloons to measure risk taking. There's a range of such tools, commonly referred to as Balloon Analogue Risk Task (also shortened to BART)[xlviii]: "The Balloon Analogue Risk Task (BART) is a computerized measure of risk taking behavior. The BART models real-world risk behavior through the conceptual frame of balancing the potential for reward versus loss."

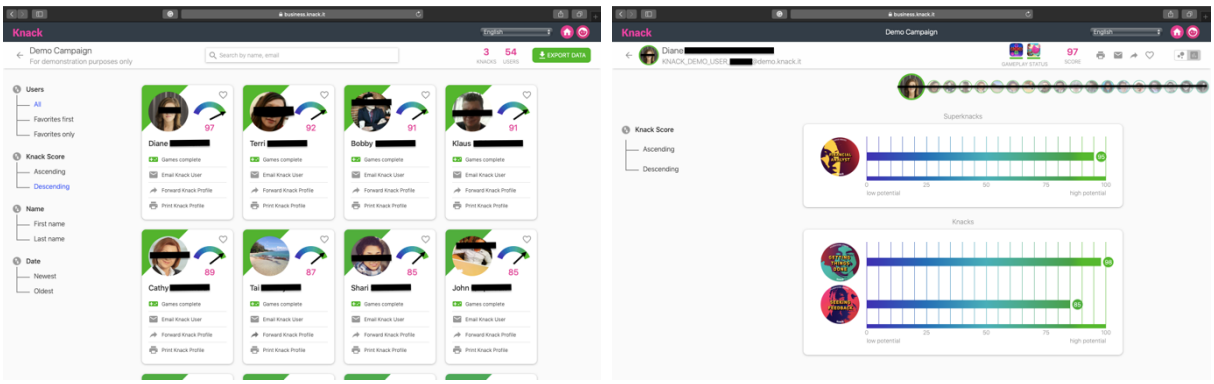


In my campaign examples, "Bomba Blitz" was generated together with "Dashi Dash" for the job of "Operations Manager", and as the sole game required to play for candidates for "Financial Analyst".

As to my overall experience from playing the three games, each of them taking 10-12 minutes to go through, I found them quite entertaining and easy to play - but hard to master. For a 41-year-old, I probably have a fairly average amount of experience in gaming: I don't play mobile games frequently, but I have tested a few, and put my share of hours into playing games in my earlier years. I would thus conditionally agree with the statement from the report "Impact Hiring: How Data Will Transform Youth Employment": "Knack's games are designed for a casual user: even inexperienced players can learn

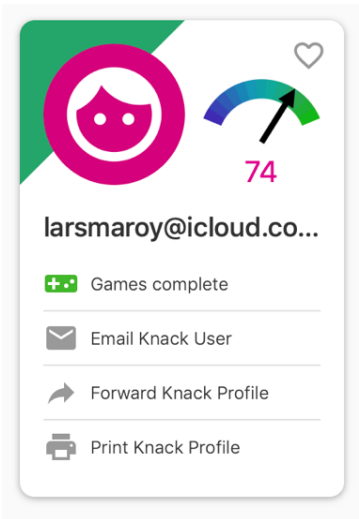
how to play the games in seconds".[xlix] If these were the first games I played in my life, or even the first games I played on a mobile device, I would probably not have scored as well as I did. It should be noted that there are no opportunities for training before the game starts and playing the game several time doesn't help either: "Knack only uses the data from the first time you play a game when providing knacks. These data are the most accurate and the most useful."[l]

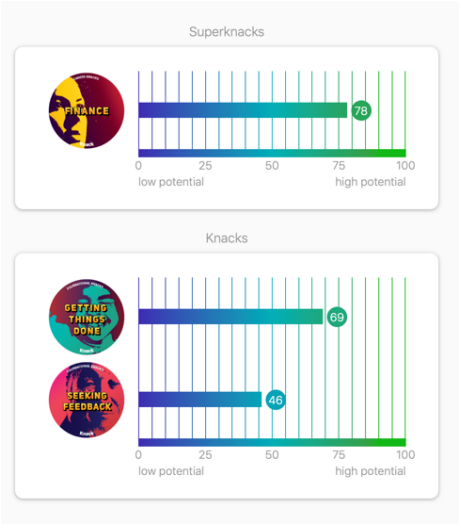
Once a new candidate has played through all the games related to the particular campaign, the employer will be notified that a new person is added to their candidate list. An example of such a list, from the pre-defined "Demo Campaign", with the specific info from the presently top ranked user (sorted by ascending "Knack Score"), follows:



As we can see, the results are here broken down to three components: "Finance", "Getting things done" and "Seeking feedback". On these, our demo user Diane X had scored respectively 99, 96 and 80, all out of a 100 percentile. This scale is explained by Knack: "A percentile score is effectively the relative rank in our worldwide reference population of over 25,000 people from 110 different countries. Example: a percentile score of 70 means that 70% of the reference population had a lower score on that knack. (Note that a lower score can sometimes be better for a particular opportunity.) »[li]

My own results from the same (demo) campaign:





PYMETRICS

New York, London, Singapore and Sydney based Pymetrics was founded in 2013. On the company web site, CEO and co-founder Frida Polli, Harvard and MIT trained academic neuroscientist and MBA, describes her personal motivation for starting the Pymetrics:

> In the day and age of Netflix, Spotify and Amazon -- platforms that take in information about you and give you personalized recommendations that seem to know you better than you know yourself – where was the equivalent for jobs? Netflix's movie recommendations are not based on their "back of the movie" blurbs. instead, they analyze movies based on deep analysis of traits and then match you based on the traits you like in movies. So why are we still evaluating people based on their "blurbs," i.e. their resumes? Why was no one applying this powerful technology to help us make one of our most important decisions – what we do with our careers?»[lii]An article in The New York Times (27.12.2017) writes about Pymetrics´ mission:
>
> "It (Pymetrics) makes software to help companies evaluate job applicants, replacing flawed methods like campus recruiting and résumé screens with a series of neuroscience-based games that are intended to be nondiscriminatory. [liii]

In 2018, Pymetrics was awarded as "Technology Pioneer" by World Economic Forum for the breakthrough work the company is doing to bring efficacy and equality to the hiring process.[liv]

As to the output from the games, the New York Times describes how they first *"are analyzed with algorithms that compare an applicant's skills with those of existing employees. The algorithm's results are then analyzed and tweaked to make sure they are not giving an advantage to applicants of any gender, race or educational background."* [lv]

Some major corporations like Unilever and Accenture (who helped fund the company through its FinTech Innovation Lab) are already using Pymetrics[lvi]. The company has raised at total of US$ 56.6 million in seven founding rounds between 2013 and 2018.[lvii]

Pymetrics' web site is targeted towards employers and candidates:

- https://www.pymetrics.com/employers/
- https://www.pymetrics.com/candidates/

Pymetrics offers a "series of behavior based games (that) collect millions of data points, objectively measuring cognitive and personality traits."[lviii]

As per the time of this writing, Pymetrics doesn't offer candidates an open channel to testing the games themselves. Unlike Knack's games, which can be tried by anyone with a smart phone, access to Pymetrics' platform goes exclusively through an employer who has entered into collaboration with the startup. In my view, this reflects the stronger emphasis that Pymetrics puts on analysing the skills of present employees prior to assessing new candidates. Not least, it reflects on their ambition to reduce bias in the recruitment process, as they require doing a "tweaking", some sort of manual processing, of the gathered data to neutralize the effect of prior biases (as per above). For the Pymetrics platform, the results of one single tested candidate has little value by itself. As assessment of a candidate is useful in relation to the assessment of the company at its present employees.

On this backdrop, I regret not being able to test fully the employer side of Pymetrics´ platform (as described in the methodology chapter). What I was able to test, was a demo

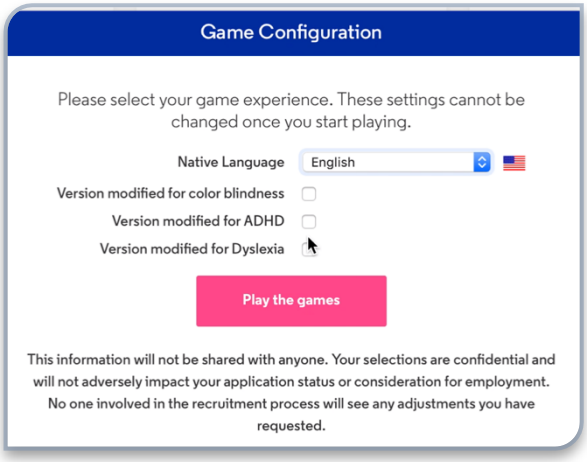
site for candidates, gaining their perspective on the process, from which I could to some extent deduct the function for the recruiting organizations.

After a straightforward registration, and reading through and accepting several privacy statements, an upbeat video with some brief instructions (1 min) followed, and then:

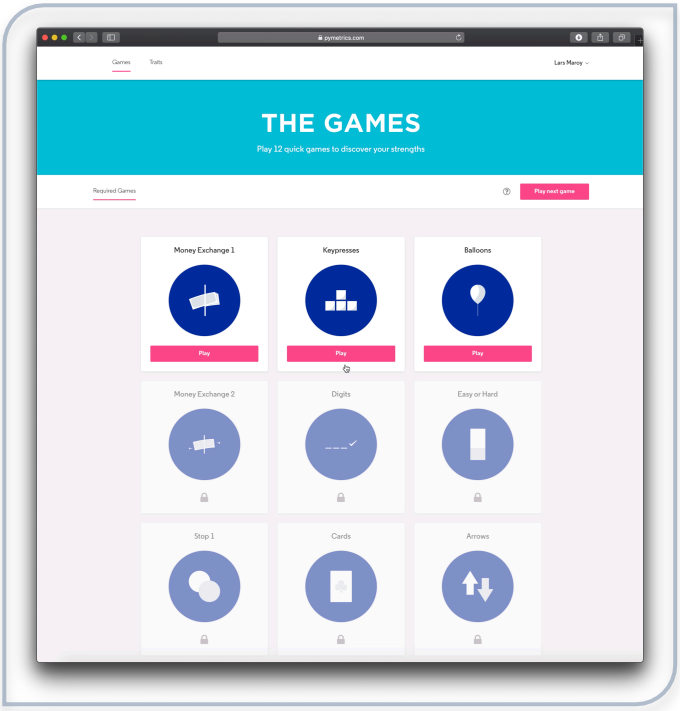
"Some advice from the experts:

- Play the game in a distraction free environment on a mobile device or a computer.
- Each game takes 1-3 minutes and don't need to be played in one sitting.
- There's no way to do "well" or "poorly" on the games."

Next step is a "Game Configuration". The system supports 19 different languages, including two versions of Chinese. None of the Scandinavian languages are presently supported.



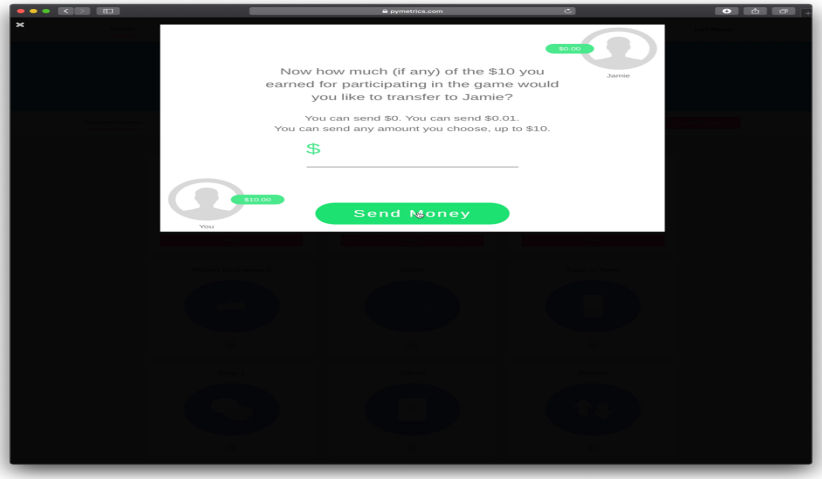
The next screen presented "12 quick games to discover your strengths". Of these, three were playable, the rest momentarily locked, but opened once I'd played the first tree:



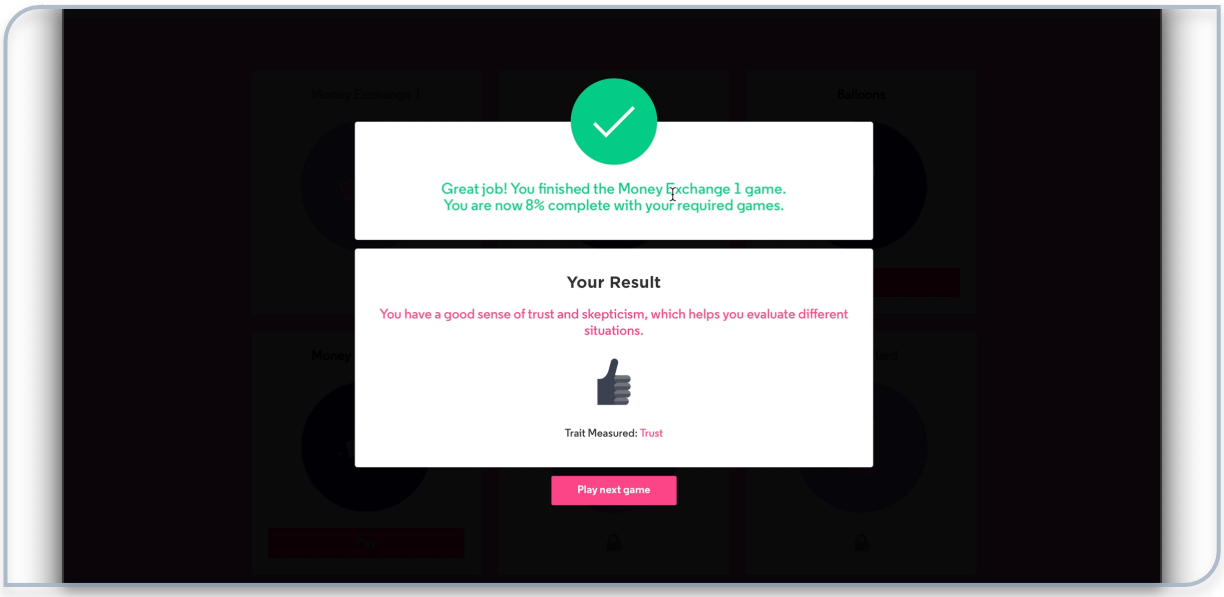
As Pymetrics explains in the FAQ section of their web site:

"The games assess 50+ different traits. We tell you all the things we assess at the end in the full report. We don't tell you which games test which traits because that would spoil the fun!"[lix]

The first game I tested was "Money Exchange 1": I was handed $10 in game money, and "paired with another random participant" (most likely the game engine itself), called Jamie, and "the opportunity to send some of the money you just received to Jamie. This $ is tripled once it's transferred to Jamie's account. Then Jamie can return all, some or none of this money back to you. (…) Whatever money you get back from Jamie is what you get to keep at the end of the day."



I sent Jamie $5, got $8.61 back from him, and then had to rate Jamie´s fairness in this transaction (on a scale from 1 to 10). This produced the following feedback:



The game "Money Exchange 1" clearly reminds me of the classical game theory case of "Prisoner´s Dilemma"[lx], and Pymetrics says it's meant to measure the trait of *trust*.

In the following, instead of going through all of Pymetrics' 12 games, I'll provide some examples from the range of games measuring different traits.

The "Keypresses game" is simply about the player pressing the spacebar on the keyboard as many times as possible in a limited period of time, measuring, according to the provider, the trait of *processing*.

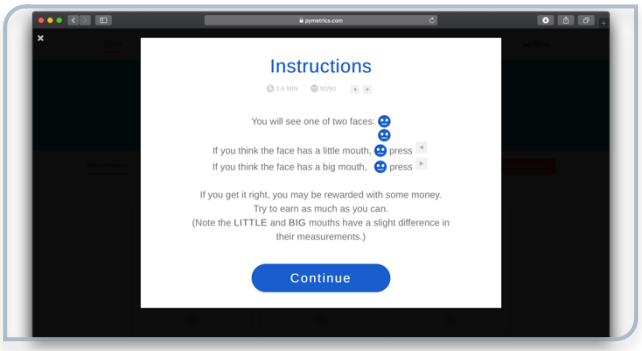
The "Balloons game" is another example from the mentioned The Balloon Analogue Risk Task (BART) genre (as per the review of "Bomba Blitz" above). You gain money for each pump on a balloon, but loose it all if you over inflate them. The balloons have different explosion points, and it's not immediately apparent which ones can take more air, although a pattern starts to emerge from the order of the balloons. The game is meant to test the trait of *risk*.

"Digits" is a game that tests the skills of *memory*. You need to remember a sequence of numbers. The length of the sequence is increased with each right answer and reduced with each wrong one. The player is encouraged to "Please do not write your answers down while playing."

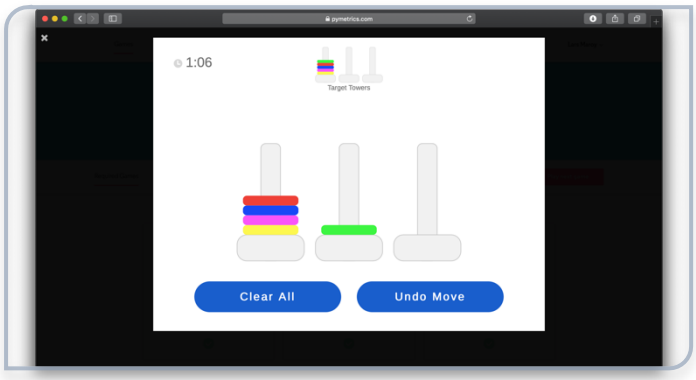
In the "Stop 1" game, you need to press the spacebar when and only when a red circle appears on the screen. The trait measured is *attention*.

"Arrows" is another version of the same, but specifically testing for the trait of *learning*.

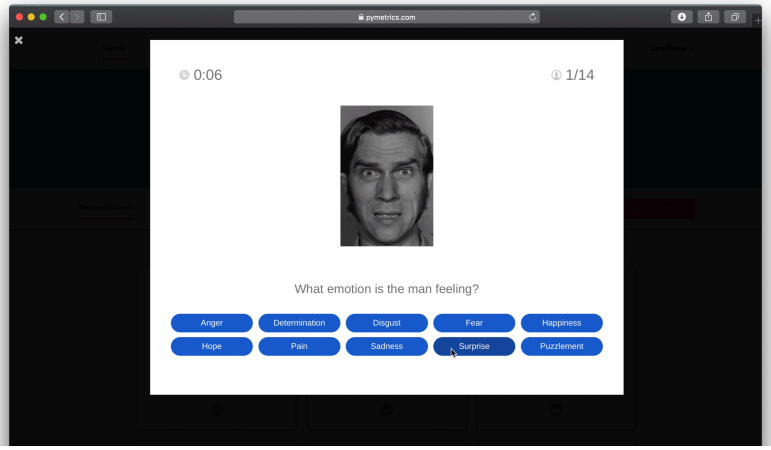
The same trait, *learning*, is measured with the "Lengths" game, where you compare the size of the mouths of two different faces, presented in different orders. The feedback I received after completing that challenge was" "You tend to be motivated by both internal and external factors."



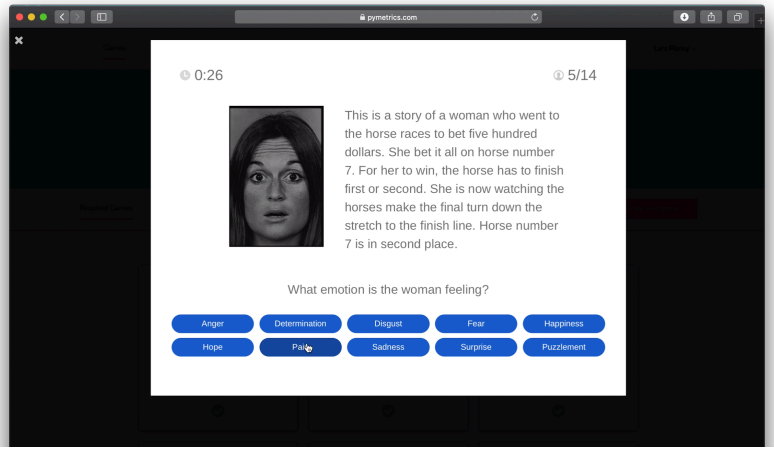
The skill of *planning* is measured by "The Towers", a variant of the classical game "Tower of Hanoi."[lxi] Moving only one disk at a time and using as few moves as possible, you need to make the bottom towers match the target towers."



The Faces game is all about *emotions*. You're shown photographs of people with different facial expressions and given seven seconds to choose the word describing what you think that the person in feeling.



Some photographs are accompanied by a short story describing a situation. You than have 30 seconds to indicate the word best matching the facial expression given the stated context.

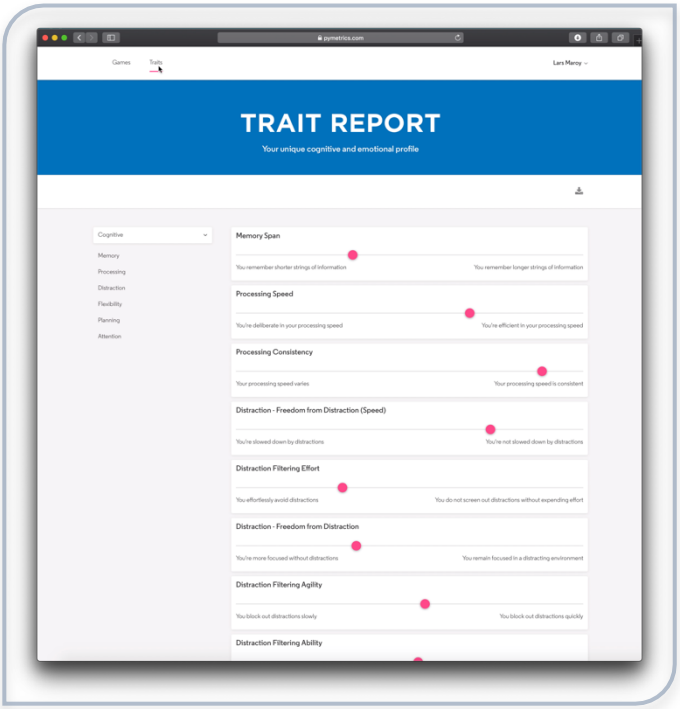


My personal feedback was: "You read other people well, but don't rely solely on facial expressions to understand the emotional context of a situation."

"Faces" clearly draws from neuropsychological research on facial expressions, not least Paul Ekman's[lxii], whose Pictures of Facial Affect (POFA)[lxiii] seems to have provided the pictures used.

Having concluded the 12 games, I was asked to "Help us get to know you better" by providing my gender, current location, ethnicity, and age. "We will anonymize and use this data to improve our user experience and our ability to remove bias from our algorithms. None of this data will be shared or affect your job matches."

At this point, I had spent 35 minutes after logging into Pymetrics´ candidate demo solution. I was now able to access my personal "Trait Report – Your unique cognitive and emotional profile." The report gave me my results, presented as indication on scales for traits such as Memory Span, Processing Speed, five different indicators for Distraction and eight related to Risk, and a range of others.



My immediate reaction from seeing the trait report was twofold: Firstly, I did somehow recognize myself in the findings. Secondly, I was unsure how to interpret the results, and how to use them. The difference from Knack´s trait presentation is obvious: Much greater detail, but fewer conclusions or advise to candidates. This probably explains why Pymetrics´ solution is still not accessible for the general public.

FOCUS GROUP WITH INDUSTRY EXPERTS

The focus group meeting took place in Bergen, Norway, on 4th of December 2018. The participants were:

- Tarald Eik Mong, Senior Executive Researcher, Mercuri Urval Norway
- Andreas de Lange, Recruitment Manager for IT in Bergen & Stavanger, Experis Norway.
- Anders Kleppe Norheim, Executive Researcher, Mercuri Urval Norway

The author opened the focus group meeting by showcasing Knack and Pymetrics as per the review above. The following is a thematically organized edit of the subsequent discussion.

WHAT'S REALLY NEW ABOUT THE NEW TOOLS?

The group agrees that the so-called "war for talent" has been raging for many years. They think that the gap between the skills that businesses in Norway and the rest of the Western World will be needing in the coming years, and the skills possessed by potential new workers for these companies, is only increasing.

> We need to invest to fill that gap and think about recruitment in new ways. That's where tools such as Pymetrics and Knack could come to use.
>
> Tarald

As this group of recruiters sees it, the novelty of Pymetrics and Knack lies not in the psychometric test components in the games themselves. These seem to be based mostly on well-established industry standard tests, such as assessments built on "Eckman´s faces", that individually are available to the market. What separates these two services from solutions already known to the group, is the way that the tests have been bundled and developed into games that are easy (or relatively easy) to start interacting with. Since the apps can be used on common appliances such as mobile phones, the tests can also be distributed broadly. The assessment of the results is automatized, and the candidates ranged according to their mental capabilities and behavioural style, in principle with little or no human interference. That's the real news, according to the focus group.

Further, the group notices the ambition of the providers to develop their solutions into *platforms*, directly connecting candidates with businesses. As they see it, such a match-making platform, a "Netflix of the recruitment world", could be particularly useful in cases where the predictors of job success are uncertain, or the skills needed to succeed in the work at hand are hard to define. In those situations, it would make sense to test the top performers in a field and compare them with the test results of candidates, with the algorithms possibly identifying other and stronger links than human recruiters would be able to.

STRENGTHS AND OPPORTUNITIES

The group agrees that, according to research, it would generally improve the prediction for future job performance to include a cognitive ability test in a recruitment process. At the present, these sorts of tests are often administered late in the processes, typically before the finale interview, when the recruiters are down to two to four candidates. Prior to that, the rest of the candidates have been screened off based on other criteria. We'll be returning to the reasons for this limited use of cognitive tests, which partly is related to perceived lack of relevance for some positions, especially on the side of the recruiting managers, and partly to time and cost.

The gamification of the tests and the automatization of the assessment in Pymetrics and Knack represent a great potential for such tools to scale the deployment of psychometric test from the employer side (since the marginal cost of the tests is very low) and to increase the use of the tests from the candidate side (since the motivation for taking a gamified test would be higher than for standard "boring" psychometric test, and the practical barrier for taking it lower, as the candidate can pick both time and place for the test, submitted over their mobile phone or PC). In sum, that could result in psychometric tests being used more broadly and earlier in the recruitment process, most likely increasing the prediction value of the overall recruitment process, at an aggregate level.

> It would clearly increase the effect of cognitive abilities tests if they were used before the client or we recruiters had made up our minds about the candidates. It's hard to change a first-impression.
>
> Anders

The group further notices that it seems easier to adapt or tweak these tests for specific positions, then is the case with today´s leading psychometric tests. According to the group, they are now almost always deployed in a one-size-fits-all approach, running identical tests on candidates for all positions – be it for a sales manager or an IT consultant. The recruiters judge the over-all performances, without differentiating between the skills relevant for the different positions, since they often don't have the time or the money to do so. As they say, maybe relational abilities were more important in one of these positions, or strategic thinking, or cooperation skills, or risk taking? Tools such as Pymetrics and Knack could more easily be custom fit, increasing their practical value.

> Customization will make these new solutions appear more relevant for of our clients than the tools we use now, hopefully making them trust their predictive value more than they do today.
>
> Anders

As the group sees them, these solutions are particularity useful for positions where the candidate base is very limited, in order to identify and attracting new talent (so-called

sourcing). With their low barriers to start playing, the fun and challenging experience the games offer, these services could bring new candidates into the loop. That could provide the recruiters with access to candidates that they'd otherwise never have found, and the candidates access to jobs they'd never have thought of. The group imagines appealing to the candidates´ competitiveness, for instance by posting a campaign link on a social media site with the tagline "Are you the best front-end developer in Bergen?" Those who get a good score would see the question: "Would it be ok if one of our recruiters gives you a call to discuss some career possibilities?" The group pictures that such an effect could generate "an extra handful of relevant candidates" in a pressured job market. As a result, they see the greatest potential for these tools in relation to the sectors that now struggle the most to find relevant and available candidates, for instance IT, consultancy, engineering and sales management.

> The struggle within these sectors is getting so desperate that our clients are ready to be salvaged by any tools that could provide them with more high potential candidates.
>
> Tarald

In the opposite situation, when there's a lot of possible candidates and the job of picking the right candidate ("kissing the right frog") is demanding (so-called *selection*), the effect of Pymetrics and Knacks could be different, but still relevant. The group discusses the imaginary case of Kiwi (a Norwegian supermarket chain) looking for a supermarket cashier, receiving 300 applicants for a position. Which of the candidates is the most intelligent, the best multi-tasker, the most emphatic, etc.? That's hard to test at scale with the tools available today, resulting in employers screening by CV and references, possibly creating a "catch 22-situation" for applicants with less relevant background info. The final decision is typically based on the impression the top candidates give when met face-to-face in an interview setting. Instead, by exposing all 300 to psychometric tests, you'll probably bring forward other types of candidates, possibly with higher potential than the ones winning the selection game today.

Another benefit from using games as a gateway to recruitment is getting to test the candidates' *flexibility*, a trait that presently is often included in job requirements. When candidates apply to a position and a game shows up, do they think "This is exciting and new, I'm willing to give it a go." Or do they say "What's this? A game? What kind of employer is this really? I don't want to work here!" The fact that the candidate is willing to play the assessment games could therefore work as a selection mechanism by itself. Some of the candidates might also develop a more positive attitude towards employers using gamified tests: "Wow, this company must really be innovative!"

> People with a lot of gaming experience will give these tests high face validity ("the transparency or relevance of a test as it appears to test participants"[lxiv]), especially Knack´s test. They´ll say "Finally, an employer who takes my skills seriously!", and be highly motivated by the tests and find them relevant.
>
> Anders

On general terms, the group agrees that tools such as Knack and Pymetrics are more apt for certain junior positions where you need to "hire for attitude, train for skills". For these positions, the employer doesn't necessarily require prior expertise in the field, since there is usually some time to learn the robes before the newly hire's job performance becomes critically important to the employer. In these situations, the group finds that these tools could work well - and that they'll predict who should be hired to be trained better than their existing methods do.

In that respect, the group believes that these kinds of tools will become more relevant in the future that they are in the present. As one of them puts it: "Knowledge comes with a rapidly decreasing expiration period". Businesses find themselves in ever-changing environments, and adaptability is kind. The knowledge you are hired because of, will not be the knowledge that's important for your employer two to three years down the line. Your CV and former experience as indicators of knowledge might thus become less relevant in the future. As these psychometric tests focus more on potential and intellect and personality, and not on knowledge and experience, that'll increase their relative value if workplaces continue to evolve in that direction. Going for potential instead of knowledge might even challenge the current educational system:

> As employers, we use the educational system not only as educators but also as gate keepers, trusting them with selecting the highest potential computer scientists, for instance. In the future, maybe large companies could find these high-potential kids themselves, train them for specific skills during a year or two and set them to work. That could provide the companies with loyal and relevantly skilled talent.
>
> Andreas

Further, the focus group member notice that both Knack and Pymetrics stress the importance of a bias-free recruitment, Pymetrics most outspokenly. The recruiters say that all (or almost all) of their clients claim that they´d like a more diverse workforce. They are therefore likely to prefer a less biased recruitment process over a more biased recruitment process.

> The problem is, most of our clients seem to forget the importance of diversity once the recruitment wheel starts turning.
>
> Tarald

The recruiters in the group agree that bias is a huge problem in recruitment. If they used a game-based test as the start of the recruitment process, versus reading all candidates' applications and CVs like they do today, that'll in effect broaden the gate for many candidates, for instance those with foreign sounding names. The recruiters could present the most fitting profiles to their clients, asking them if they're willing to take a look at the top candidates, regardless of background and experience, 'checking out wild cards' as they say.

> These solutions can help us challenge our clients: If they mean what they say about increasing diversity, are they willing to these solutions that could produce a less biased outcome than the tools they´re familiar with? I wonder if they're willing to go that far. And it's not easy being the first recruiter to confront a paying client with that approach.
>
> Anders

WEAKNESSES AND THREATS

Knack and Pymetrics are gamified psychometric tests. Research clearly indicates that psychometric tests should be given weight in recruitment processes. But stakeholders at the companies looking for candidates don't necessarily see it that way:

> In my experience, the employers usually don't care about the results from the cognitive test.
>
> Andreas

Andreas presents the group with two fresh examples from his work as a recruiter to illuminate his point of view: His clients were two different companies in the IT sector. In both cases, they were down to the final two candidates for an open position, and cognitive tests were deployed. One of the candidates scored significantly better than the other in both examples but was still turned away in favour of the one with the lower score, due to other factors: "Charisma" in one of the cases; "A better culture fit" in the other.

This leads the group to discuss a possible challenge for Knack and Pymetrics and other providers of psychometric tests at scale: If employers don't trust cognitive tests in general, finding them irrelevant for the job performance they're after, how would they feel about putting a lot more emphasis on psychometric tests, like the use Knack and Pymetrics would imply? The employers would first need to be convinced about the research on the predictive value of different test forms. They would, for instance, need to agree that references are weak predictors for job success, while employers typically give them a lot of weight today.

To avoid biases sneaking into the process, they would also need to decide a priori which factors are important for each position, and not rationalize as they frequently do today, changing the weight they place on the different criteria when they hold the different candidates up against the criteria set. Summing up, the group is not overly optimistic as to changing their clients' view on recruitment methods, in the short run.

> We recruiters need to go in the front here. The clients trust us, and they would, at least to some degree, believe us if we said that these are the tools you should use for this job.
>
> Tarald

The group agrees that it would take some convincing to get the recruitment industry on board as well. To be convinced themselves about the true potential for using psychometric tests as a first gateway to a broad range of jobs, the group would like to see more research on their predictive force:

> Even though cognitive tests are the strongest predictors at an aggregate level, there's still a lot of variance left, with many other factors also bearing a predictive value. Also, it's likely that cognitive abilities predict better for certain types of jobs then for others, and I would think, until I see proof to the contrary, that the same goes for these game-based tests.
>
> Anders

Another challenge to game-based services is that some candidates are likely to find such tests to be more stressful than the present alternatives, and would need other settings, such as a written application letter or a one-to-one chat or a case-based test, is the recruiters are to see their true potential. Further, the group fears that the face validity for these solutions decrease with the candidates' rising age. The result could be exclusion of older candidates.

> Some older people even struggle to use more than the basic functions on their smart phones. It's naïve to assume that they will take to these games like duck to water.
>
> Tarald

The act of playing a game will in itself feel less natural for an average 60 y/o then for an average 20 y/o, who is likely to have far more experience in apps and games. If you're not used to gaming, you might lack the self confidence in your own abilities to perform well in such a situation, compared to more traditional "pen and paper" tests. Especially when introducing a time element, like these games do. The clock is ticking, and you probably understand that your score is influenced by the time needed to do the tests.

> Candidates unfamiliar with games could be excluded. That's bad for them of course. However, this might actually be the selection that many employers today actually want!
>
> Anders

The group appreciates the potential for these platforms to reduce the effect of biases and heuristics in the recruitment process. But they fear that they could end up discriminating against other traits instead. Older people, and those unfamiliar with games, have already been mentioned. Another example could be a candidate with ADHD. Some of them could score poorly on the games that require the player to stay focused on one task for a period of time. That would be a relevant screening effect for a job that required focus, but not all jobs do. You could in such cases miss out on some good candidates. The

recruiters underscore that these new tools, as all others, need to be applied with caution and preferably by skilled practitioners in the field.

> Also for these reasons, I'd be more open to using this kind of tools for junior positions.
>
> Andreas

Junior positions are also typically the ones with the highest number of applicants. In these processes, Knack and Pymetrics could serve well for selection. However, the group sees a danger in that sort of testing. They fear a "winners-take-it-all situation", where the top 5-6 % of candidates are the ones that all employers fight to get, even for low-impact positions such as the supermarket cashier mentioned above, because the candidates have the potential to succeed in all kinds of work. They would be selected again and again, jumping from one position to the next. The system would work well for these top workers, and the individual companies using them, but just in the short run, before they're off to the next employer.

> I find this quote relevant: "Hard work beats talent when talent doesn't work hard." In tests such as these, you'll find the most talented candidates. But it's not certain that they'll put in as many hours in the job they get, then someone less talented but with more grit, who's always had to fight for it. I fear that these tests and using them early in the process will skew the selection towards the more intelligent candidates, possibly at the expense of the most hard-working ones.
>
> Andreas

The group members emphasize that screening is not the specialty field of neither one of them. They are mostly occupied with the process of *sourcing*, the identification of new talent that the companies in Norway´s present job market are most willing to pay for. As for using Knack and Pymetrics' for sourcing, the group says that this "hire for potential, train for skills approach" is great in theory, but challenging in practice. In almost all of their current missions, the clients want candidates with relevant experience, who can be found quickly, and hit the ground running, performing well from day one. They're referred to as "plug-and-play candidates". As a result, recruiters are forced to screen for knowledge and background and not only for potential.

> For instance, if I'm looking for a full-stack programmer, it doesn't help if a candidate performs very well on these game tests, if they don't know the first thing about programming. Those candidates would have to be identified based on more traditional methods than gamified psychometrics. Someone working at MacDonald's won't be considered for a data scientist position, unless they actually know about data science.
>
> Andreas

The main problem for the recruiters and their clients today is the lack of relevant applicants to many high-impact positions. Their clients pay them typically 150.000 -

200.000 NOK to find their "dream candidate". Advertising on social media, traditional press or job search sites yield few relevant results. That is where the recruiters provide their greatest value, in the form of a wide network. More often than not, they claim, the best candidate appears through their personal connections. Our recruiters are divided in the question of whether putting a game such as Knack on a platform such as Facebook would increase the supply of relevant candidates for a senior position such as full-stack developer.

> You'd have to believe that the most relevant senior candidates are positive to playing a game to get a job. I'm not so sure.
>
> Andreas

The focus group explains how many of the most attractive candidates are tired of being approach by recruiters.

> I actually think that many of the top candidates won't play the game, even if they think it looks like fun, simply because it might trigger a call from one of us.
>
> Anders

Reluctance to consider job changes is not a trait exclusive to IT specialists and engineers. In surveys, people have ranged job seeking as one of their least favourite activities:

> Surveys have shown that candidates would rather pull a tooth than start a new job seeking process.
>
> Tarald

Tarald shows the group a recent survey[lxv] that states that 60 % of Norwegians aged 18-67 find the process of getting a new job "very stressful" or "quite stressful", actually scoring above pulling a tooth when it comes to stressfulness.

That could be an issue for recruitment services dependant on broad outreach. Even though most people are open to changing their job if it's for the right one, according to the group members, games-based tests connected to candidate-to-business platforms are likely to generate most offers for positions that are not the candidates´ dream jobs. If job seekers are convinced to play a game and a recruiter calls you with a proposal that you don't find interesting, will you then play again, or more likely delete your profile from the service to avoid future irrelevant offers? For these platforms to work, the candidates need to be more open to new types of possibilities, positions and sectors than they are now.

FURTHER DEVELOPMENT NEEDED

Which further development would our group of recruiters like to see in Knack and Pymetrics and similar solutions to increase their utility as process tools? The first thing they mention, is the need for a **Norwegian language** version. The instructions, especially in the Pymetrics test, seem complicated enough even for a native English speaker, and could easily discriminate others. The group assumes that a Scandinavian version is underway, and hope for that version to be as good as the original, which they find not always to be the case with traduced recruitment tools; Something tends to get lost in the translation, between languages and between cultures.

Another issue is **privacy**. These tests generate a large amount of highly sensitive data on a large number of candidates. The group notices that the providers state compliance with GDPR. And all candidates need to accept the legal terms before playing, but these are most likely related to the laws and regulations in the test providers´ home country, the U.S. In order for the recruiters to use the tools in Norway, they would need to be harmonized with privacy and employment and anti-discriminatory laws in Norway. Knowing that legislation usually moves slower than technology, that could lead to some difficulties for tech optimistic recruiters wanting to adapt cutting edge solutions.

Validity and cheating are also brought up. The incentives for cheating are clearly present in any candidate assessment. With tools that only make sense if the recruiter trusts their candidate scores, is it acceptable let the candidates take the tests in un-controlled environments? The group members had diverging views on that question.

> If these tests are taken at home, I'd be too tempting to hand over the phone to a friend who's a brilliant auditor, if it could help to bring me into a process for an auditor position. We're only human, after all.
>
> Tarald

Other group members argue that forcing the candidates into the office would ruin the positive effect of bringing in new candidates due to low entry barriers.

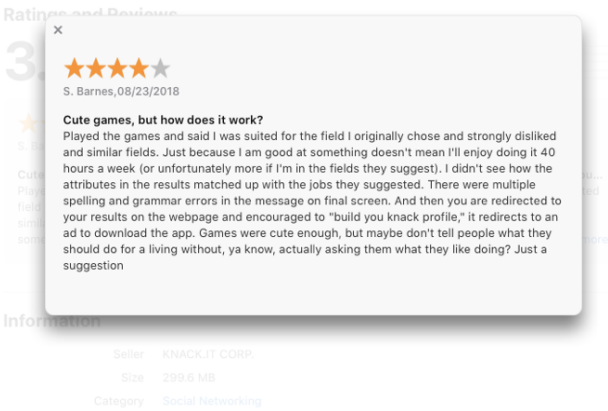
The group also sees a potential in game-based psychometric tests being developed further to include more directly relevant **job-related simulations.** Situational judgement tests predict job performance well. If they could be tailor-made and introduced into these games, for instance simulating a sales situation when recruiting a sales person, the recruiters and not least their clients would be forced to give these types of assessments more weight.

To prevent the "grasshopper effect" with top scoring candidates jumping from position to position, the group wonders if **loyalty** as a trait could be measured as well. To their experience, most employers would rather hire a candidate in the 70th or 80th percentiles, who stays on for seven years and shows great progress during the period, than someone in the 90th percentile, who is more likely to jump fence a lot sooner.

Another issue debated is whether or not Knack and Pymetrics should ask the candidates about **fields of interest,** in order to improve the relevance of job offers:

> It’s easier for recruiters to call someone for a position in finance if they’ve already stated an interest in that sector.
>
> Anders

I showed the group a review of Knack from App Store that highlighted Anders´ point of view:



> Playes the games and said I was suited for the field I originally chose and strongly disliked and similar fields. Just because I am good at something doesn’t mean I’ll enjoy doing it 40 hours a week. (…) maybe don’t tell people what they should do for a living without, ya know, actually asking them what they like doing?
>
> App Store reviewer

The downside to including such a question would be a more limited scope, going against the spirit of these solutions as the group sees them, introducing once again a bias from the side of the candidate as to potential positions. A large part of the benefit from these recruitment games is putting new ideas into the candidates´ heads. To the recruiters’ experience, job seekers generally have very little imagination for possible career paths, often thinking very traditionally. To take out the full potential of Knack and Pymetrics and similar solutions, a more open and bias free approach is needed from the employers as well as from the candidates.

CONCLUSION: GAME-CHANGERS OR EXTRA TOOLS?

To conclude, I asked the focus group to assess the disruptive potential of services such as Knack and Pymetrics on the recruitment business: Are they "game changers", like Knack proposes on its web site? The recruiters agree that both tools could represent a change to recruitment, but over time:

> I don't believe these gamified tests will set off a revolution in our industry anytime soon. But they might help us increase the predictability of the recruitment process on the future job performance. I don't yet see them as pure substitutes to the tools we use today.
>
> Tarald

The most relevant positive effect on the daily work of recruiters would come from Knack and Pymetrics used as sourcing tools, bringing new candidates into a tight labour market. They also see them as effective selection tools for junior positions. They're more sceptical when it comes to senior positions, fearing that they might turn away candidates unwilling to play a game to change jobs. As for executive recruiting, the group doubt these tools will ever be able to compete with the old face-to-face and "who's in your contact list" approach.

Tarald gave an example where his company once recruited an oil engineer to a cancer research institute, based on the fact that the mathematical formulas that are used to estimate where there's oil in a reservoir are very similar to the mathematical formulas used to estimate how breast cancer spreads through humane tissue. The institute would normally never have considered hiring someone with such a "wrong CV match". And the candidate would never have thought of entering into health care. The group finds in very interesting to see if these kinds of tools could automate and scale that type of innovative search for candidates and linking them to new kinds of possibilities.

> I see a huge potential there. Using meteorologists for big data analysis, instead of fighting over the too few trained computer scientists. But the most interesting connections are the ones that we humans fail to see, but that an AI could, given enough input and feedback.
>
> Tarald

Still, the focus group members agree there are limits to potential as a sole guiding principle. If you´re looking for a programmer, it doesn't help if a candidate has the potential of becoming an amazing programmer, if they don't know the first thing about coding. Companies are not ready to take over the role of colleges and universities, at least not yet.

Concluding, the group expresses some scepticism towards the new type of tools, largely taking the role of "devil's advocates" on behalf of their clients. Per now, they see Knack and Pymetrics as new tools for recruiters adding to the ones in their existing toolboxes, more than complete game changers. But they´re open to reconsidering, after seeing the tools traduced to Norwegian, tested thoroughly by recruiters, and generating a representative selection of candidates. If that pool has an interesting profile, the group thinks that their clients and thus themselves are ready to be persuaded by gamified psychometrics.

CHAPTER 5: DISCUSSION

INCREASED PULLING FORCE

In a two-tier job market with increased competition amongst employers to attract the most qualified applicants (Michaels *et al.*, 2001 in Searle, 2009). The recruiters in the focus group spend their days in the top level. Their main headache over the last years has been the lack of available specialised talent for high-impact positions. That resonates with a survey reporting that 65 percent of recruiters claim that talent shortage is their biggest hiring challenge[lxvi]. Conversely, problem and possibility are interlinked: The ability the recruiters have to unearth and bring to the table such hard-to-find, tough-to-get candidates is precisely what their clients pay them to do – it is the main value proposition of the recruiters.

How could game-based tests, such as those provided by Knack and Pymetrics, come into play in the context of recruitment? According to the literature on gamification (e.g., Deterding, 2012, and Aparicio *et al.,* 2012), the use of game-elements has the potential of rising motivation and engagement. In our case, that would imply that presenting the invitation to apply for a position by means of a game instead of using an ad or similar static medium, would increase the likelihood that people who already have a job, will present themselves as candidates for a new one. I agree with the focus group members that "Play a game and receive career recommendations" is a proposition that could evoke the interest of many a passive job seeker. If so, that would be no small achievement, knowing from surveys (quoted by a member of the focus group) that people would rather pull a tooth than start applying for a new job.

The big question is: Would the candidates attracted through a game be the ones the recruiter wants, and the hiring organization needs? The overall utility of a recruitment/selection system should be assessed by the quality of the applicants attracted (Carlson et al., 2002). How Knack and Pymetrics will fare in that sense remains to be seen, as I will return to later.

BETTER SELECTION

The above was related to the recruitment, or the attraction or sourcing, part of the hiring process. Next step is selection, choosing the best candidate to fill an open position. The core of Knack and Pymetrics are gamified psychometric tests, purporting to gauge in a bias-free manner the candidates´ potential for different careers, by assessing their cognitive, social and personality. Amongst HR scholars (e.g., Searle, 2009) there appears to be consensus on cognitive ability being, on average, the most reliable predictor for future job performance. The focus group members are aware of this research. Still, they seldom use the ability tests until the finale heat of a recruitment run, with only a small handful of candidates remaining, the rest eliminated on other grounds than on their cognitive abilities.

When asked about why they do not use cognitive tests more often and earlier in the process, the focus group participants said that their clients, managers at the recruiting organizations, distrust them. They are not unaware of the scientific findings but seem to ignore them in real life, a well-known effect in HR literature (Rynes *et al.,* 2007, in

Searle, 2009). Instead, they appear to choose their tools based on ease and familiarity (Bersin and Chamorro-Premuzic, 2019).

This underuse of cognitive tests is not necessarily due to HR professionals distrusting them on average. My guess is, they would probably recommend other HR professionals to use such tests for a majority of positions. Their scepticism is just related to the position that a given recruiting manager is focusing on at the moment, with the particular needs of that position. Since it is time-consuming and costly to adapt the current cognitive tests for the specific requirements of a given job position (as per the focus group interview), the recruiters almost always administer them in a one-size-fits-all model. As a result, the HR managers don't find the tests relevant for their specific needs (low face value). As a result, they lean on other selection mechanisms instead. That would typically be a job interview or a reference check, both weaker predictors of job performance on a group level (Searle, 2009). With the gamified assessments offered by Pymetrics and Knack, and the latter in particular, could provide an efficient way of tailoring the ability tests to the specifics of a job opening, increasing their (perceived) relevance for any given position.

Lack of face value was not the only reason given by the recruiters as to the underuse of cognitive tests. The cost and time-consumption related to the administering and assessment of the present-day tests was also mentioned. In that respect, the game-based the solutions are promising in the sense that they automate this labour-intensive process. Ability tests can therefore be included in more recruitment processes, and more candidates can be tested in each of them. Given the lower psychological barriers for taking a game-based versus traditional ability tests (e.g., Lappalainen, 2017), chances are, more candidates are also willing to take them.

Based on the above, I agree with the focus group that introducing game-based tools could be beneficial to the processes of recruitment and selection. "Extra tools in our toolbox," as they put it. Related but not 100 % comparable, the vast majority of recruiters in a survey are favourable to the use of artificial intelligence, with 96 % of them believing that AI can greatly enhance talent acquisition and retention.[lxvii]

However, how much of a real *game changer* for recruitment and selection could gamification become?

GAME CHANGERS?

There group seemed to agree on the view that gamified tests are not setting off *"a revolution in our industry anytime soon,"* as one of them put it*.* However, I'm convinced that Pymetrics and Knack (and other providers of game-based hiring solutions) are presently sharpening their knives, preparing for that exact revolution. These actors not in the market of adding just another tool to the toolbox of recruiters. They´re here to cut themselves a nice slice of the US$ 200 billion recruitment industry, estimated to grow to US$ 334 billion by 2025. They will try to do so, not by providing stand-alone game-based assessment tools, but by turning these tools into fully automated match-making platforms. Knack puts it straight: "Knack is first and foremost a platform, not simply an assessment tool. »[lxviii] While Pymetrics talks of "Netflix, Spotify and Amazon -- platforms that take in information about you and give you personalized recommendations"[lxix]. I

doubt that Pymetrics would've been able to raise US$ 56.6 million if it were not for their ambition of becoming "the Netflix-like recommendation algorithm for jobs."[lxx]

Knowing how industry after industry has been turned upside-down by the introduction of high-tech newcomers, one could have expected the industry experts in the focus group to have been more alert to the disruptive potential of Knack and Pymetrics and similar, known and unknown, providers. When they seem not to be so overly preoccupied, the focus group members act as described by practice and theory on disruption (by Clayton Christensen, in particular[lxxi]): They look at the new entrants and rightfully conclude that the products and services they offer don't off measure up to the quality of the products and services they offer themselves. The value propositions of Pymetrics and Knack don't presently match the needs of Mercuri Urval´s or Experis´ customers. These clients are at the higher tiers of the market for recruitment services and they would presently rather pay a human recruiter NOK 300k to find a proven full-stack developer than a digital platform NOK 1k to find one with the *potential* for becoming a great one.

The problem for the recruitment industry incumbents, as the focus group members represent, is that the entrants don't need to attack from the top. Knack and Pymetrics (and *their* rivals, such as Google for Jobs[lxxii]) are much more likely to leave the most demanding and sophisticated customers to the traditional recruiters, and instead focus on the bottom of the market – for starters. "An innovation that is disruptive allows a whole new population of consumers at the bottom of a market access to a product or service that was historically only accessible to consumers with a lot of money or a lot of skill. » (Christensen, 2019)[lxxiii]

In the following I'll attempt to outline how the disruption case could potentially play out:

DISRUPTING FROM BELOW

In this context, I'll be using the term *disruption* in the "orthodox" sense[lxxiv], whereby a smaller company with fewer resources, in our case Knack or Pymetrics, can successfully challenge established incumbent businesses, in our case Experis or Mercuri Urval. These companies are now focusing on improving their products and services for their most demanding and profitable customers, typically helping them fill high-stakes well-paying tech jobs.

The recruiters undoubtedly give value to their clients. Yet, it might be that they provide too much value in some parts of the process, like the general cognitive tests, that today´s hiring managers don't seem to trust and therefore might have little actual willingness to pay for. At the same time, these search and selection companies are not as actively selling themselves to potential clients looking to fill "low-stakes" positions, like for a supermarket cashier, or medium-range positions such as a middle manager for the same supermarket. The incumbents claim, probably rightfully so, that there´s too little willingness to pay for their services in that market. There´s also typically an element of prestige on the side of the established companies, adding to the barrier they feel towards moving down-market.

High-tech entrants, like Pymetrics and Knack, could target those overlooked segments at the bottom of the market, thus gaining a foothold in the recruitment industry by delivering the basic functionality that the general manager of the supermarket is looking

for, at a reasonable price. The entrants could then move upwards in the market, aiming for junior positions, avoiding at first the current booming industries, like IT and engineering, where the willingness to pay is high, inversely proportional to the number of available candidates for any position.

Gaining momentum, and showcasing their success, Knack and its likes could start courting the incumbents' mainstream customers, delivering the performance they require, while focusing on their unique advantages that drove their success in the low market; broader attraction, automated selection, and so on. They might still be aiming for the junior positions that are too unprofitable for the low-tech recruiters, who are now chasing higher profitability in even more-demanding segments, encouraged by clients who are sceptical to new tools and reluctant to leave any decisions to robots.

Eventually, some of the recruiting managers will start testing these game-based recruitment and selection tools on also senior positions. Incumbents classically respond by moving even further upmarket instead on fighting vigorously downwards. As they are fighting harder for the lucrative *executive* search and selection, their traditional mainstream customers are starting to accept the entrants' high-tech offerings in large volumes. *Classical disruption has occurred*. Not necessarily because the newcomers provided a service that was qualitatively equal or better in every sense, but because they targeted a neglected customer base with a "good enough" low-cost product that matched these customers´ basic expectations for a recruitment service – and built on that.

I am not claiming to know how the future of recruitment will appear. But I would not be surprised if it were to develop in a pattern such as the mentioned. I am also not saying that Pymetrics or Knack will be the entrants disrupting the recruitment industry. It might be another startup, or very well one of the tech giants such as Google. Regardless who eventually comes knocking, how should the incumbents react? My main piece of advice would be to prepare early. They should acquire a real option[lxxv], by expanding into AI fuelled game-based assessments. That could be done by internal development or by procuring a start-up. Such a move would help them resist competitors attempting to disrupt their businesses from below, as they could pre-emptively offer low-cost solutions to clients at the bottom half of the market for recruitment and selection services. Seeing the situation from our Norwegian or Nordic point of view, one could contemplate the potential of one provider taking the possibly very lucrative position of "Finn.no for recruitment" (Finn.no is the dominant classified advertisements website in Norwegian, and the largest website in the country in the number of page views[lxxvi]). Such a monopolistic scenario would depend on this being a "winner takes it all-market." That would again depend on the actual market potential for gamified assessment platforms and not least on how network effects will play out.

GAMIFIED ASSESSMENT PLATFORMS

Could these solutions take the step from process tools to platforms? The recruiters in the focus group talked about using Knack and similar solutions to build themselves a larger pool of candidates out of which to pick "wild-cards" for some of their clients´ open positions. However, what if the game-based assessment providers made that pool by themselves?

A platform owning the data coming from the candidate assessments (versus strictly renting its capabilities to paying clients), could compare the data from one candidate to the data from others, thus providing the candidates with advice based on real careers (like Knack to some extent seem to do already). There could be a feedback from the candidates about the perceived relevance of the advice, tuning the recommendation given in the future (machine learning). That would create a network effect since the value to any one of candidates using the service would increase with the total number of candidates using it. However, that effect will level off after a certain number of candidates, making it relatively easy for other competitors, "the next Knacks," to enter the market. The network effect alone does not lay the ground for a monopolistic situation.

The "winner takes it all" comparison with Finn.no would only be relevant from the moment the gamified candidate assessment platforms turn to recruiting organizations, offering to connect them directly to potential candidates. The more candidates who are filing their profiles on the platform and registering their traits by playing the games, the more recruiting organizations would want access to that platform to fill their vacancies. Once candidates from a platform´s pool are hired into a company, the platform could gain insight into the performance of different trait bearers in different positions, tweaking their match-making algorithms further. That would increase the value for businesses and thus their demand for the platform, which again will increase the value for potential applicants, in a virtuous cycle. Welcome to the most frequent wet dream of tech entrepreneurs: to own a platform fuelled by *two-sided* network effects (Eisenmann *et al.,* 2006)[lxxvii]

If such a double network effect happens, will we then necessarily see a winner-takes-it-all market, with one dominant platform, or could there still be room for several suppliers? Using more than one platform could be costly or unpractical to businesses, which per definition would favour a single dominator in each job market (like Norway). However, it's not unlikely that the market will be segmented, with specialized platforms for different segments, such as IT, health care, education, and more. That would go against the very spirit of the game-based solutions I've studied, which is all about sector neutrality. Yet, sectorisation could still be relevant in order to increase trustworthiness towards recruiting managers who are likely to still think traditionally about their jobs - possibly until they one day are replaced by a millennial hired through a game-based platform.

In order for any of these effects to happen, some key elements must be in place. Some of these challenges are systemic, external to the control of the game-based candidate assessment providers. Others are more basic requirements, hygiene factors if you'd like, that need to be solved internally at the providers before they are likely to experience a broad market impact.

SYSTEMIC CHALLENGES TO GAME-BASED ASSESSMENT PLATFORMS

ALGORITHMIC AVERSION

If psychometric assessments are applied at scale by the use of gamification, and that leads to better hires (as per the introduction), then rational and enlightened recruiters and hiring managers would gladly hand the reins over to the recruitment platforms and their algorithms. According to the focus group, however, it is not quite that easy. First, both recruiters and hiring managers are likely to be reluctant to give up power to *choose*, a fundamental human reaction. Second, gamification is something new, and "new things are scary," as one of the recruiters put it. More fundamentally, there might be an element of "algorithmic aversion" (Dietworst. 2015)[lxxviii], which could over-shadow the positive effect of the data-based selection: "Even when faced with evidence that an algorithm will deliver better results than human judgment, we consistently choose to follow our own minds.» (Michelman, 2019)[lxxix]

Algorithmic aversion would be a peril to game-based assessments. Therefore, when analysing the future potential for such tools, you would want to consider whether or not the trust in algorithms is increasing or decreasing. It would be great news for both Knack and Pymetrics if the trust in algorithms is growing. Possibly it is already stronger than many scholars believe, as a recent HBS study indicates (Logg, 2018)[lxxx]. A possible explanation for such a development could be that we are increasingly exposed to the advice of robots (e.g. in correcting our writing, directing us while driving, guiding us to the cheapest online retailer, matching us with a date). Why would the managers of the future, a majority of which will soon be millennials, unknown to a world without Google, smartphones and social media, not trust an algorithm's recommendation for the highest potential candidate if they trust an algorithm's recommendation for a life companion?

As to the opinion of the general public, there is still a way to go in order to gain the acceptance of algorithmic hiring: In a recent study (Pew Research Center, 2018)[lxxxi], 57 % of U.S. adults say that automated resume screening of job applicants is unacceptable. 67 % of say that automated video analysis of job interview is unacceptable. They are concerned that such programs violate privacy, are *unfair*, *remove the human element from important decisions,* and worry that *"humans are complex, and these systems are incapable of capturing nuance."* Tech optimists, and I count myself as one of them, tend to think that all others think alike, which they, of course, don't. At least some of Knack and Pymetrics´ potential clients are likely to worry about the public acceptance of their hiring tools.

POTENTIAL OVER SKILLS

At the core of both Pymetrics and Knack is a belief in the value of people´s potential vs. their manifested knowledge and background, opening up for fresh candidates. To which extent does this approach resonate with the current wants and needs of real-life recruiters.?

The focus group was concerned that the emphasis on candidate potential could go against the current needs of their highest-paying customers. Indeed, there are too few relevant candidates, and that calls for an approach focused on potential. Yet, the same clients are typically also in a hurry. They hire these recruiters in order is to find "the

perfect candidate," who could "hit the ground running," as they said, and perform well from the start. For that reason, they would not recommend someone working at MacDonald's for the job of a programmer, if that person did not master any programming languages, regardless of that person´s cognitive capabilities per se.

In a New York Times article[lxxxii], I found an example from a company that, on the face of it, seemed more open towards potential. Thanks to an AI-driven recruiting service called Eightfold.ai, the resumé from a prospective data scientist was highlighted, without that candidate having any apparent training or experience with data science. However, he did have a master's degree in statistics and knew several computer programming languages, which are relevant for a data science position. The selection was not all about abilities and traits, even though it' claimed that the abilities and traits made the candidate stand out from the rest.

Interestingly, on Pymetrics' help desk, someone has asked a question that relates specifically to the question of potential versus experience:

> "Will I be competing with people of all ages and experience for the same jobs?"
>
> Pymetrics' answer: "Companies generally compare candidates with similar experience and education level for a particular position so you will be evaluated with a pool of candidates who have backgrounds relevant to the job requirements.[lxxxiii]

It would seem that even the providers of these services are sceptical to the relative importance of potential, or – more likely, they have adjusted their services to accommodate the view of hiring managers. I find that to be a sensible, pragmatic approach on the side of the providers, even if it contradicts with their fundamental trust in potential.

The same goes for the candidates´ fields of interest. As one of the recruiters in the group said: It´s easier to approach someone with an offer for a job in a given sector if you know that the person has expressed an interest in that line of work. You don't necessarily like doing something, just because you have a good head for it.

Apparently, both Knack and Pymetrics could benefit from including questions about the candidates' background (possibly automated by harvesting data from the candidates´ LinkedIn profiles) and of interest fields. Such a tactic could also defend the platforms from critical remarks such as this one, related to gamified candidate behaviour tests:

> In one way, the experience strips down the idea of choosing a job to what it is: a capitalistic matchmaking of labor and employer. Fitting the optimal cogs in the optimal slots. That razor efficiency might pay little attention -- if one even has the privilege to seek it -- for what you ***want*** to do, perhaps out of a sense of social duty or personal meaning.[lxxxiv]

INCREASED EFFORT IN ON-THE-JOB TRAINING

To the focus group´s experience, there's seldom time anymore for the in-house development of new hires. That resonates with studies showing that provisions of on-the-job training have fallen substantially, down approx. 50 % in the US and the UK over the last two decades[lxxxv]. Lack of available staff training is not good news for tools designed to put true potential over proven qualifications. Luckily for providers such as Knack and Pymetrics, change seems to be on its way. Potentially the beginning of a new trend, spend on staff training at American companies rose by almost a third from 2016 to 2018.[lxxxvi]

Rapid technological change reducing the value of "old" skills is one reason why companies are again starting to put money into training their employees: "Employee skills stay relevant for only three years" according to Diane Gherson, head of HR at IBM[lxxxvii]. Skill shortages in the candidate market, a concern repeated several times during the focus group conference, is another reason for the rebirth of in-house training. Both of these are good news for those advocating for the "hire for potential, train for skills" approach. The recruiters say they find it increasingly hard to find "the perfect candidate", or in some cases, any qualified candidate at all. Possibly, their clients understand that the skills hired for today are going to be less relevant in a few years. That would incentivise them to look for applicants´ potential, and then provide adequate training (not just once they're hired, but most likely continuously during their careers, if the technological development keeps up its pace) in order to bring out that potential.

The developments outlined above, concerning algorithmic aversion, the need to emphasise potential over knowledge, and related to the latter, increased effort in employee training, in sum would indicate that that utility of tools like Knack and Pymetrics is likely to increase in the future, as the focus group industry experts also concluded.

BASIC REQUIREMENTS

Apart from the above-mentioned changes to society or business ecosystems, that seem to align with the game-based solutions' unique selling points, there are some basic requirements related to the specifics of the solutions themselves that I believe must be in place before a wide adoption will take place:

The solutions must feel immediately relevant to candidates. The major effect is hopefully the appearance of a career possibility. The first feedback will, however, always be the results from the game-based assessments themselves, in the form of a trait report (Pymetrics) or "my knacks" or career advises (Knack). In order to benefit from word-of-mouth, augmenting the network effect, that immediate feedback to candidates is vital. I believe both providers have a way to go in that respect. Personally, I found Knack´s analysis to be just partially interesting. My top strengths were "Being careful", "Getting things done" and "Acting with self-confidence." I would've preferred if the system quantified the findings and showed relative results on the candidate side like it does on the employer side: For instance, "Congratulations! You are better at getting things done than 62% of Norwegians". That would've been an incentive for bragging on social media, like Knack encourages users to do, spreading the word about the platform.

As for Pymetrics, I found the trait report challenging to interpret. According to the report, I'm more thoughtful than efficient and more trusting than sceptical. Interesting, but overwhelming as these were two of the scores on a total of 40 different traits. I that sense, Pymetrics could learn something from Knacks´ more simplified presentation. The candidates must feel they receive something of immediate value for their time spent playing the games (1-3 of them in Knack´s case, 12 in Pymetrics´).

The user experience is paramount. These tools will only work as promised if they provide a low barrier to candidates, bringing in applicants that would not materialize using the incumbents' recruitment techniques. Thanks to gamification, motivation, and engagement could increase substantially (Deterding, 2012). It is hard to imagine these games competing with all sorts of casual games in the commercial market, on the game experience alone. Likely, they do not need to, given the extra value added (as per the last point).

However, a hygiene factor for a game used in recruitment is a flawless user interface. We have come to expect games to work correctly, with little patience for loading and little acceptance for flaws, for instance due to lack of adaptability to different screen sizes. I find that the two services, and Knackapp in particular, still have some way to go on the user experience side of the games. As an App Store reviewer of Knack commented (albeit as far back as in September 2015, and I do see that the solution has improved drastically since then):

> This app and the two games I played associated with this app (Wasabi Waiter and Balloon Brigade) are so full of bugs. (...) And what makes it even worse is the fact that these aren't just games. Employers are using the data uploaded to aid them in the hiring process. If you thought you got mad when something messes up that's just a game, imagine how lad you would be if it effected your ability to get the job you're applying for!!!

The user experience is especially important as the game providers need to prove that the solutions **also work for non-gamers**. Even if a majority of the population (in the Western world) play games occasionally or often (Nielsen, 2018), a large minority still doesn't. Even though the focus group participant could be right in his assumption that the very act of screening for gaming skills could be relevant for many employers, all those non-gaming hiring managers out there will need to be assured that the results from the gamified solutions are neutral to the candidates´ experience with games. Putting it simply, they need to believe that *they* could've been hired using Knack or Pymetrics, that these tools are not only for millennials or younger. After all, the workforce is aging, and it will continue to age, because of low birth rate and more people working for more years.[lxxxviii]

Authentication: If the clients are to trust the results from the tests enough to give them decisive weight in the following process, there should be methods for ensuring that the test taker is the real candidate, taking the test unaided (Bartram, 2000). As of now, I can see now such control mechanisms in the platforms themselves. The control would then need to be done by the recruiters or the hiring organizations.

Control over practice: As a hiring manager, I would also have liked to know if all test takers have had a similar amount of practice for the test form (Bartram, 2000). Both

Knack and Pymetrics claim to have reduced the issue of "control over practice" by allowing only the results from the first play-through of a game to count towards your final score. (Example: [lxxxix]) Yet, some sceptics would argue that the issue has not been completely dealt with, as some candidates could be suspected of wanting to improve their score and subsequent change of getting hired by testing the games using a secondary email account, before submitting a "real" test. As one of the focus group representatives put it: The candidates are only human, and the agent-principal problem[xc] in hiring is substantial.

Proven bias-free: Famously, in 2018 Amazon.com stopped developing a recruiting algorithm after learning that it had been unfairly penalizing female candidates. The reason why bias had sneaked in was that the company's history of hires in the male-dominated tech industry had taught the system that male attributes were preferred.[xci] The company received a substantial amount of criticism, but also some praise: "What makes Amazon unusual is that it actually did its due diligence, discovered the troubling bias and decided against using the algorithm.»[xcii]

In the mentioned Pew Research Center study, 58 % of Americans claim to feel that computer programs will always reflect some level of human bias – although 40 % think these programs can be designed in a way that is bias-free. The providers of gamified candidate tests need to prove they are. "Plausible deniability»[xciii], pretending to not know about biased algorithms to avoid responsibility in the face of future lawsuits, won't do it for the companies using such services, and the providers need to reassure them. The fact that the alternative, keep basing the hiring process on human judgement, is likely even worse concerning bias, will not help either. A possible parable can be found in self-driving cars: A 2018 study[xciv] from the Society for Risk Analysis found that the public will not accept self-driving technology before it is shown to be approximately four to five times as safe as human-driven vehicles. It might be the same with algorithmic hiring. More independent research is needed to prove that game-based candidate assessments reduce to a great extent the impact of biases on the recruitment process.

Accountability: Regardless if the game-based tools are used to hire someone or to fire them, how do you explain why person 1 was a *yes*, and person 2 a *no*? I´ve not studied the technical side of nor Knack or Pymetrics. I do guess that, since there is machine learning involved, there will be an element of "black-box": "The black box metaphor dates back to the early days of cybernetics and behaviourism, and typically refers to a system for which we can only observe the inputs and outputs, but not the internal workings.»[xcv]

Decisions deriving from a system that´s unexplainable per definition, since its designed to identify patterns and react, and not to "reason" in the human sense of the word, are likely to attract criticism, like in the case of the babysitter assessment service Predictim:

> (…) critics say Predictim and similar systems present their own dangers by making automated and possibly life-altering decisions virtually unchecked.
>
> The systems depend on black-box algorithms that give little detail about how they reduced the complexities of a person's inner life into a calculation of virtue or harm. And even as Predictim's technology influences parents' thinking, it remains entirely unproven, largely unexplained and vulnerable to quiet biases over how an appropriate babysitter should share, look and speak.[xcvi]

Predictim was using AI to analyse the social media profile of candidates for babysitting jobs. Some candidates that received something less than a perfect score, and therefore didn't get hired, wanted to know why, and there was no easy way of extracting that explanation from the system. The same could happen with applicants using Knack and Pymetrics. More importantly, assessing the market potential of the solutions: Hiring managers could imagine that candidates who were turned down might be wanting an explanation, and possible they could demand one, according to the laws in some countries. The providers need to deal with the demand for accountability, providing solutions that can explain the decisions made.

Local adoptions: Needles to mention, the solutions must be adapted to the particular needs of the different markets the entrants would like to approach, for instance translating to Norwegian language and harmonizing with our national privacy and labour laws, and more. An example relating to the last point: EU´s general data protection regulation article 22[xcvii] states that «the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her». Google and Facebook have solved this issue by allowing their customers for instance to opt-out of personalised recommendations. A game-based solution presented as the entry gate to working for a company would face other legal challenges that would have to be dealt with.

To conclude, the respective above-mentioned basic requirements are not trivial to fulfil, and in total they constitute quite a challenge for any small or medium sized newcomer in the market for game-based candidate assessments. Since the first supplier who is able to scale rapidly enough to handle these issues will have a considerable advantage over the rest, odds are increasing for a winner-takes-it-all situation. Pymetrics´ seven rounds of funding might indicate they have come to the same conclusion. For the same reasons, consolidation amongst candidate platform providers is also likely to take place.

LIMITATIONS AND FURTHER RESEARCH

The present thesis has evolved into a case of possible classical disruption, with the focus group members representing scholarly examples of incumbents, seemingly ignorant to the real potential of entrants. Therein lies also the study´s major methodical weakness: How certain am I that the recruiters in the group are representative for the recruitment industry as a whole, enabling me to generalize and claim that the industry has a bias that prevents them from seeing the full potential of new technology? My selection of focus group participants all work in the high-end market for recruitment services and they are likely to have been coloured by colleagues and clients at that higher tier.

It could be that other recruiters would be more directly open to the benefits of the newcomers´ solutions. It is also possible that hiring managers, who typically hire for low-end positions themselves, would have responded differently to the new tools and their potential. Therefore, a study with a more diverse selection of industry sources could've made the conclusions stronger. Possibly, a larger, quantitative study with recruiters and hiring managers would've generated a more representative outcome.

In the same context, I should add that none of the final members of the focus group are directly involved with the strategic operations of the companies at which they work. The discussions in the group are likely to have turned out differently had one of the executives that were candidates to join, been able to participate.

Furthermore, gamification is a relatively new field of study, with the sub-field of gamification of HR functions in general and of recruitment and selection in particular, even younger and less mature. A study of how the use of gamification for personnel attraction and selection changes as the field matures would be interesting to see. Such a study could also showcase the heterogeneity of providers of game-based solutions – a topic on which I only touched briefly since the particularities of the different providers were of less relevance to my research question.

Lastly, although not the focus of the present study, it would strengthen the case for these game-based psychometric assessments if there were studies validating them directly to on-the-job performance, instead of to other general intelligence instruments (Kline, 1993 in Searle, 2009), like we´ve seen in the study by Lappalainen *et al*. Lacking such a source, the discussion had to be based on face value, or else, the assumption that these tools actually work as the providers claim.

CHAPTER 6: CONCLUSION

Answering the first part of the research question, "Which are the potential benefits and limitations of gamified psychometrics for recruitment and selection?", I would like to focus on three positive and one negative aspect:

Based on my findings, the main benefit from using game-based assessments in the *higher tiers* of the job market is their potential for sourcing in new, relevant candidates, of which there seldom are enough.

In the *lower tiers* of the market, the solutions could partly automate the time-consuming work of selection, resulting in higher cost-efficiency.

In both ends of the market, turning psychometric tests into games could increase their utilization, by reducing the barriers for delivering as well as for taking such tests. Knowing for instance that cognitive ability is a strong predictor for job performance at an aggregate level, using game-based assessments could lead to better hires in general.

Among the obstacles facing the providers of game-based candidate assessments, I would probably highlight *algorithmic aversion* as the one most critical to handle. These tools will only be allowed to work if people trust their decisions. In that context, Amazon´s 2018 AI-recruitment failure did not do their colleagues at Knack and Pymetrics any good. A fair amount of bulletproof independent research and good PR is needed to convince both the public and hiring managers about the advantage of algorithmic recruiters.

Concerning "How likely is the prospect of game-based assessment providers disrupting the recruitment industry?", my conclusion is that such disruption is likely to take place, in one form or another. The recruitment industry is of considerable size ($200 billion) and still quite low-tech, and it would go against all business trends in the last decade if it were to stay that way for much longer. The industry is particularly vulnerable to the intrusion of high-tech entrants such as Pymetrics and Knack if it is, in fact, blind to the true potential of new technology. I have found indications that the recruitment industry is only partially aware of the dangers of disruption, with some limitations concerning the selection of sources to the study potentially weakening that conclusion.

CITATIONS

[i] Michaels, E., Handfield-Jones, H. and Axelrod, B., 2001. *The war for talent*. Harvard Business Press.

[ii] http://gates-solutions.com/blog/the-war-for-talent-is-over-the-talents-won-the-theory-of-the-high-street/ accessed 30.01.2019

[iii] 2017 Deloitte Global Human Capital Trends https://www2.deloitte.com/content/dam/Deloitte/global/Documents/About-Deloitte/central-europe/ce-global-human-capital-trends.pdf accessed 30.01.2019

[iv] https://www.wsj.com/articles/the-loneliest-job-in-a-tight-labor-market-11545224400 accessed 25.01.2019

[v] Searle, R., 2009. Recruitment and selection. *Human Resource Management, A Critical Approach*, pp.151-168.

[vi] Carlson *et al.,* 2002, in Lievens, F. and Chapman, D., 2010. Recruitment and selection. *The SAGE handbook of human resource management*, pp.135-154.

[vii] Barber, A.E., 1998. *Recruiting employees: Individual and organizational perspectives* (Vol. 8). Sage Publications.

viii Deterding, S., 2012. Gamification: designing for motivation. *interactions*, *19*(4), pp.14-17.

ix Bartram, D. (2000). "Internet recruitment and selection: Kissing frogs to find princes." International Journal of Selection and Assessment 8(4): 261-274.

x Lawrence, 1999 in Bartram, D., 2000. Internet recruitment and selection: Kissing frogs to find princes. *International journal of selection and assessment*, *8*(4), pp.261-274.

[xi] GLOBAL RECRUITING TRENDS 2018 by LinkedIn, https://news.linkedin.com/2018/1/global-recruiting-trends-2018 accessed 18.01.2019

[xii] https://www.theguardian.com/lifeandstyle/2015/nov/22/why-job-interviews-are-pointless accessed 15.01.2019

[xiii] https://www.forbes.com/sites/forbeshumanresourcescouncil/2018/02/12/job-interviews-dont-work-how-to-rethink-your-interview-process/#13e641c732e8 accessed 15.01.2019

[xiv] https://joshbersin.com/2018/11/ai-comes-to-recruiting-will-interviews-go-the-way-of-the-dinosaur/ accessed 15.01.2019

[xv] Morgeson, F. P., et al. (2007). "Reconsidering the use of personality tests in personnel selection contexts." Personnel Psychology **60**(3): 683-729.

[xvi] White, T. L., et al. (2008). "Test-Retest Characteristics of the Balloon Analogue Risk Task (BART)." Experimental and Clinical Psychopharmacology **16**(6): 565-570.

[xvii] https://www.hindawi.com/journals/ijcgt/2011/282345/ accessed 25.01.2019

[xviii] Bersin and Chamorro-Premuzic in MIT Sloan Management Review, winter 2019, p. 7

[xix] https://www.theguardian.com/careers/handwriting-analysis , accessed 13.01.2019

[xx] http://www.flume.com.br/pdf/Scmidt_The_validity_and_utility.pdf p. 265, accessed 13.01.2019

[xxi] Deterding, S., Dixon, D., Khaled, R. and Nacke, L., 2011, September. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15). ACM.

[xxii] NIELSEN GAMES, GAMES 360 U.S. REPORT 2018 https://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2018-reports/games-360-2018.pdf accessed 13.01.2019

[xxiii] https://www.port.ac.uk/study/postgraduate/postgraduate-research/research-degrees/phd/explore-our-projects/gamification-of-society accessed 25.01.2019

[xxiv] Deterding, S., 2012. Gamification: designing for motivation. *interactions*, *19*(4), pp.14-17.

[xxv] Hamari, J., Koivisto, J. and Sarsa, H., 2014, January. Does gamification work?--a literature review of empirical studies on gamification. In *2014 47th Hawaii international conference on system sciences (HICSS)* (pp. 3025-3034). IEEE.

[xxvi] Lowman, G.H., 2016. Moving beyond identification: Using gamification to attract and retain talent. *Industrial and Organizational Psychology*, *9*(3), pp.677-682.

[xxvii] Armstrong, M.B., Landers, R.N. and Collmus, A.B., 2016. Gamifying recruitment, selection, training, and performance management: Game-thinking in human resource management. In *Emerging research and trends in gamification* (pp. 140-165). IGI Global.

[xxviii] Nieborg, David B. "America's Army: More Than a Game.", 2004

[xxix] Lopez, J., 2011 in Chow, S. and Chapman, D., 2013, October. Gamifying the employee recruitment process. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (pp. 91-94). ACM.

[xxx] Joy, M.M., An investigation into gamification as a tool for enhancing recruitment process, 2017

[xxxi] Lappalainen, G. and Fanni, A., 2017. *Game-Based Assessments of Cognitive Ability: Validity and Effects on Adverse Impact through Perceived Stereotype Threat, Test-Taking Motivation and Anxiety* (Doctoral dissertation, University of Calgary).

[xxxii] Methodically Breaking the Glass Ceiling, white-paper by Pymetrics (2017) https://www.pymetrics.com/science/ accessed 12.01.2019

[xxxiii] pymetrics: Using Neuroscience and Data Science to Revolutionize Talent Management, white-paper by Pymetrics (2017) https://www.pymetrics.com/science/ accessed 12.01.2019

[xxxiv] Narayanan, D, Gertner-Samet, A., Malter Cohen, M., Polli, F., Workforce Solutions **Review**, September 2016, p. 33

[xxxv] Agrawal, A., Gans, J. and Goldfarb, A., 2018. *Prediction Machines: The simple economics of artificial intelligence*. Harvard Business Press.

[xxxvi] https://medium.com/datadriveninvestor/mirror-mirror-whos-the-best-candidate-of-them-all-59b118e4bbb2 accessed 25.11.2018

[xxxvii] https://medium.com/datadriveninvestor/mirror-mirror-whos-the-best-candidate-of-them-all-59b118e4bbb2 accessed 25.11.2018

[xxxviii] Yin, R. K. (2016). Qualitative research from start to finish. New York, Guilford Press.

[xxxix] Rogers, E. M. (2003). Diffusion of innovations. New York, Free Press.

[xl] https://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapid=250015979 26.01.2019

[xli] https://www.rockefellerfoundation.org/people/guy-halfteck/ accessed 26.01.2019

[xlii] https://www.knackapp.com accessed 24.11.2018

[xliv] https://www.fastcompany.com/3014016/can-this-company-actually-quantify-human-potential accessed 04.01.2019

[xlv] https://www.knackapp.com/faq/ accessed 04.01.2019

[xlvi] https://en.wikipedia.org/wiki/Porteus_Maze_test accessed 04.01.2019

[xlvii] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4026685/ accessed 26.01.2019

[xlviii] http://www.impulsivity.org/measurement/BART accessed 04.01.2019

[xlix] http://www.incandescent.com/impacthiring.pdf accessed 04.01.2019

[l] https://www.knackapp.com/faq/ accessed 04.01.2019

[li] https://www.knackapp.com/faq/ accessed 04.01.2019

[lii] https://www.pymetrics.com/about/ accessed 05.01.2019

[liii] https://www.nytimes.com/2017/12/27/technology/tech-roundup-2017.html?_r=1 accessed 05.01.2019

[liv] https://www.businesswire.com/news/home/20180621005278/en/pymetrics-Awarded-Technology-Pioneer-World-Economic-Forum accessed 26.01.2019

[lv] https://www.nytimes.com/2017/12/27/technology/tech-roundup-2017.html?_r=1 accessed 05.01.2019

[lvi] https://www.businesswire.com/news/home/20180927005147/en/pymetrics-Raises-40-Million-Series-Funding-Led accessed 05.01.2019

[lvii] https://www.crunchbase.com/organization/pymetrics#section-overview accessed 27.01.2019

[lviii] https://www.pymetrics.com/candidates/ accessed 05.01.2019

[lix] https://pymetrics.zendesk.com/hc/en-us/articles/211083303-What-are-the-games-assessing- accessed 05.01.2019

[lx] https://plato.stanford.edu/entries/prisoner-dilemma/ accessed 27.01.2019

[lxi] https://en.wikipedia.org/wiki/Tower_of_Hanoi accessed 05.01.2019

[lxii] https://en.wikipedia.org/wiki/Paul_Ekman accessed 05.01.2019

[lxiii] https://www.paulekman.com/product/pictures-of-facial-affect-pofa/ accessed 05.01.2019

[lxiv] *Holden, Ronald B. (2010) in* Weiner, I.B. and Craighead, W.E., 2010. *The Corsini encyclopedia of psychology* (Vol. 4). John Wiley & Sons.

[lxv]https://finnspirasjon.finn.no/uploads/img/default/0001/01/494ceb0073489e2be536f6fff82b3129e56caf8d.pdf accessed 29.12.2018

[lxvi] http://web.jobvite.com/FY18_Website_2018RecruiterNation_LP.html accessed 27.01.2018

[lxvii] https://www.forbes.com/sites/joshbersin/2017/03/09/robotics-ai-and-cognitive-computing-are-changing-organizations-even-faster-than-we-thought/#2749ed4ca3f4 accessed 27.01.2018

[lxviii] https://www.knackapp.com/faq/ accessed 27.01.2019

[lxix] https://www.pymetrics.com/about/ accessed 27.01.2019

[lxx] https://www.pymetrics.com/science/ accessed 27.01.2019

[lxxi] http://www.claytonchristensen.com/key-concepts/ accessed 27.01.2019

[lxxii] https://www.forbes.com/sites/joshbersin/2017/05/26/google-for-jobs-potential-to-disrupt-the-200-billion-recruiting-industry/#542f9c854d1f accessed 27.01.2019

[lxxiii] http://www.claytonchristensen.com/key-concepts/ accessed 27.01.2019

[lxxiv] https://hbr.org/2015/12/what-is-disruptive-innovation accessed 19.01.2019

[lxxv] https://www.investopedia.com/terms/r/realoption.asp accessed 27.01.2019

[lxxvi] https://en.wikipedia.org/wiki/Finn.no accessed 27.01.2019

[lxxvii] Eisenmann, T., Parker, G. and Van Alstyne, M.W., 2006. Strategies for two-sided markets. *Harvard business review*, *84*(10), p.92.

[lxxviii] Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114.

[lxxix] Michelman, P.: "When People Don't Trust Algorithms», interview **with Dietvorst, B. J.,** MIT Sloan Management Review Fall 2017 Issue, https://sloanreview.mit.edu/article/when-people-dont-trust-algorithms/ accessed 19.01.2019.

[lxxx] Logg, J., Minson, J. and Moore, D.A., 2018. Algorithm Appreciation: People Prefer Algorithmic To Human Judgment.

[lxxxi] http://www.pewinternet.org/2018/11/16/public-attitudes-toward-computer-algorithms/ accessed 19.01.2019

[lxxxii] https://www.nytimes.com/2018/12/06/business/economy/artificial-intelligence-hiring.html accessed 27.01.2019

[lxxxiii] https://pymetrics.zendesk.com/hc/en-us/articles/209468926-Will-I-be-competing-with-people-of-all-ages-and-experience-for-the-same-jobs- accessed 06.01.2019

[lxxxiv] https://www.engadget.com/2018/05/04/pymetrics-gamified-recruitment-behavioral-tests/?guccounter=1 27.01.2019

[lxxxv] "Bartleby: Time to get in training", The Economist, 11.08.2018, p. 50.

[lxxxvi] Training magazine survey quoted in The Economist, 11.08.2018, p. 50.

[lxxxvii] "Bartleby: Time to get in training", The Economist, 11.08.2018, p. 50.

[lxxxviii] https://www2.deloitte.com/insights/us/en/deloitte-review/issue-21/meet-the-us-workforce-of-the-future.html accessed 19.01.2019

[lxxxix] https://www.knackapp.com/faq/ accessed 18.01.2019

[xc] https://en.wikipedia.org/wiki/Principal–agent_problem accessed 30.01.2019

[xci] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axioslogin&stream=top accessed 19.01.2019

[xcii] https://www.bloomberg.com/opinion/articles/2018-10-16/amazon-s-gender-biased-algorithm-is-not-alone 19.01.2019

[xciii] https://www.bloomberg.com/opinion/articles/2018-10-16/amazon-s-gender-biased-algorithm-is-not-alone accessed 19.01.2019

[xciv] https://www.insurancejournal.com/news/national/2018/05/31/490751.htm accessed 19.01.2019

[xcv] https://towardsdatascience.com/the-black-box-metaphor-in-machine-learning-4e57a3a1d2b0 accessed 19.01.2019

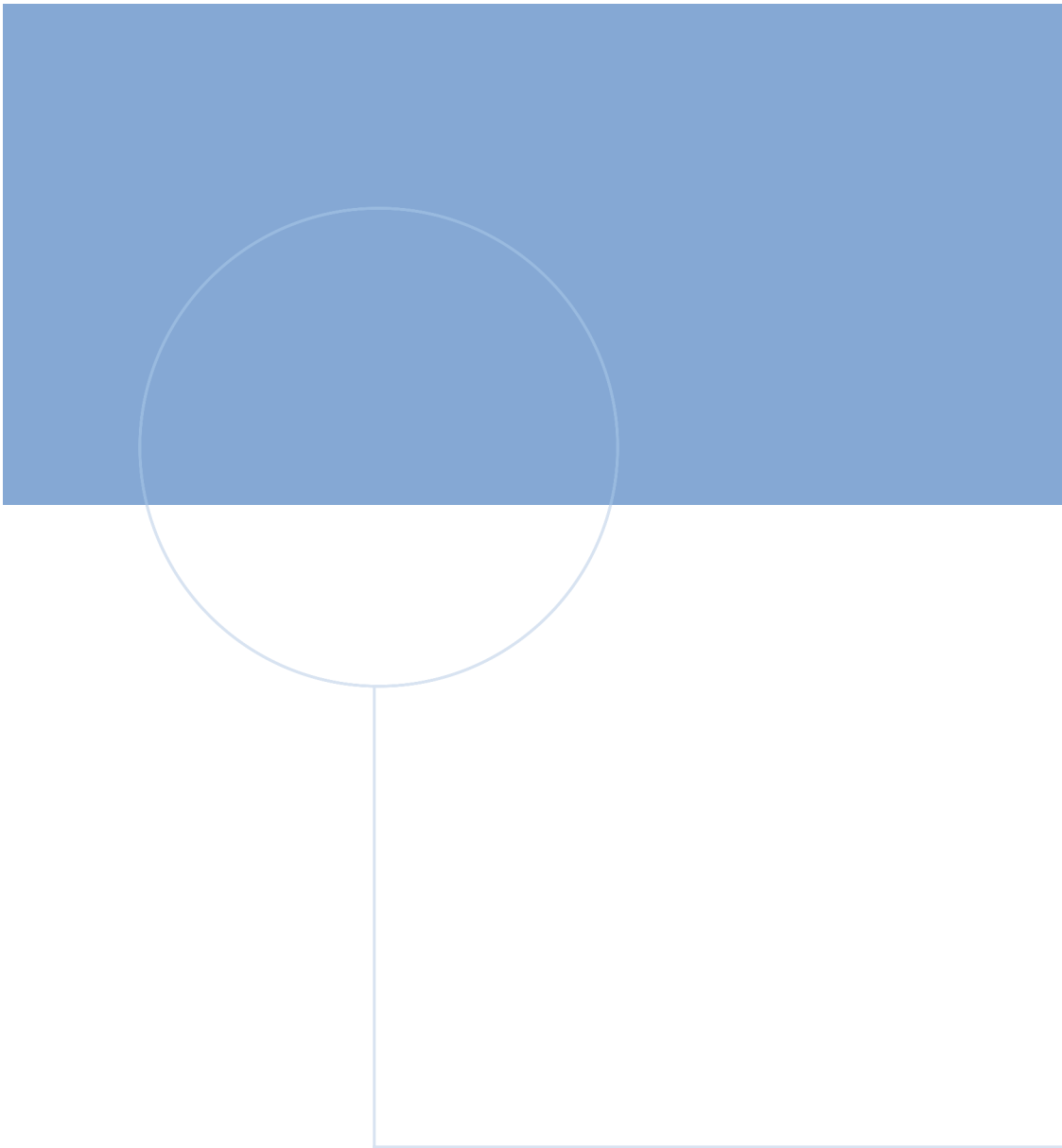
[xcvi] https://www.washingtonpost.com/technology/2018/11/16/wanted-perfect-babysitter-must-pass-ai-scan-respect-attitude/?noredirect=on&utm_term=.9a0e747a3e6e accessed 19.01.2019

[xcvii] http://www.privacy-regulation.eu/en/article-22-automated-individual-decision-making-including-profiling-GDPR.htm accessed 30.01.2019

ATTACHMENT

FOCUS GROUP INTERVIEW GUIDE

- Presentation of cases (PowerPoint, Knack and Pymetrics web sites, YouTube-video by Knack)
- Advantages over traditional tools for
    - Recruiters
    - Companies
    - Candidates
- Limitations
    - Recruiters
    - Companies
    - Candidates
- Face validity
- Result in better hires?
- Expect different reactions from a young to an older candidate?
    - Do the solutions in fact test for gaming skills?
- Bias: How important is it to companies? How could these tools come in?
- Needs for further developments?
- Potential to change the whole recruitment ecosystem, or just another tool?



NTNU
Norwegian University of
Science and Technology