Breaking Tester Stereotypes: who is testing and why it matters

Isabel Evans
Dept. of Computer Information Systems
Faculty of ICT
University of Malta
Msida, Malta
isabel.evans.17@um.edu.mt

Chris Porter
Dept. of Computer Information Systems
Faculty of ICT
University of Malta
Msida, Malta
chris.porter@um.edu.mt

Mark Micallef
Dept. of Computer Science
Faculty of ICT
University of Malta
Msida, Malta
mark.micallef@um.edu.mt

Testing is essential for successful delivery of software solutions and is not always performed by specialist testers. In earlier studies, we noted a wide diversity in the backgrounds of the testers participating in these studies, prompting our research question: Who is Testing? We conducted a qualitative study of over 70 industry testers, covering testers from multiple countries and domains, with information about their backgrounds, hobbies, roles and characteristics. We show testers are from a wide range of backgrounds, with differing needs, characteristics, and problem-solving preferences. Their roles in software projects are multi-faceted, requiring a high cognitive skill level. We discuss how to break stereotyping and best support diversity in testers' backgrounds and characteristics. We consider whether software testers are different from other software practitioners, and how understanding tester personas helps support of testers and testing.

Software Testing, People in Testing, Human-Computer Interaction, Diversity

1. INTRODUCTION

Testing is an essential part of software development and is performed by developers, specialist testers, customers, end users and other project stakeholders. The goals of testing include providing information about risk, identifying defects in the software, suggesting improvements to products and processes, and demonstrating whether products are fit for purpose (ISO (2013); ISTQB (2023)). This qualitative study of over 70 testers provides empirical evidence of testers' diverse backgrounds and preferences. Diversity breaking the stereotype of IT workers is a strength, and affects hiring practices, team dynamics, and team composition. Testers have multi-faceted, cognitively challenging roles, encouraging an increasingly diverse group joining the profession.

Using open questions, we ran a survey to improve our understanding of people testing, their characteristics, context, goals, activities, and interests. We asked '**Who is Testing?**', taking an interpretivist approach (Baltes and Ralph 2022), based on reported backgrounds and experiences of people testing software. Our findings show indicative patterns, which can inform the understanding of the testing community and potentially improve recruitment, on-boarding, training, work practices and tool support for testers. This may inform an understanding of whether testers differ from other IT workers. We wanted to let testers tell their stories, allowing emerging patterns to provide a focus on backgrounds, hobbies and roles and their relationship with stereotyping, (McChesney et al. 2022; van Tilburg et al. 2022). This is important for the recruitment and support of a wide diversity[1] of testers who represent the end users of software, and who problem solve in a multitude of ways. Breaking stereotypes and encouraging heterogeneity enables fresh perspectives and better testing of software and systems (Shneiderman 2022; Capretz et al. 2015; Robert D. Austin and Gary P. Pisano 2017).

We provide the background to our work, grounded in HCI and Software Testing, before introducing our research methodology and results. We discuss why taking a people-centric HCI lens to the testers matters for society, for building a thriving test community, and for better work practices and tools.

2. BACKGROUND

Four main concepts inform this study. We discuss software testing, a discipline of activities overlapping both software engineering and the business domain. We examine the people testing software and

their challenging but sometimes misunderstood role. We consider HCI concepts that help us understand people and their contexts. Finally, we examine stereotyping and diversity as ways of (mis)understanding people, with an impact on software engineering and society.

2.1. Software Testing

Software testing is an essential part of any software development or maintenance project. It is heavily relied on by teams and organisations to provide information for decisions about the readiness of software for its customers. Yet it is difficult and expensive (Capgemini (2022)). Definitions of testing may be divided into those that emphasise the creative, cognitive, human-focused, for example Bach and Bolton (2016), and those that sound more mechanical such as the earlier work of Beizer (1984). The former lend themselves to testing organised as people-centric with tool support, while the latter take a tools-centric, automation-driven approach. Software testing can be defined by its purposes, by its activities, or both. The purposes of testing include providing information about software quality and risk, uncovering defects and identifying conformance to requirements. Both ISO (2013) and Bach and Bolton (2016) mention test activities including planning and preparation, test execution, results analysis, risk analysis, and reviewing artefacts.

2.2. Testers: People with a Challenging Role

Over the history of the IT industry, software testing has been done both as a specialist role, and as an activity that is part of a team-wide responsibility (Tulonen (2013); MABL (2021)). Over the last few years, with a move to DevOps, the emphasis has been more towards integrating testing into a whole-team activity; for example, the State of Testing Report from Practitest (2022) states: 'In many organizations other team members also take part in the formal testing process. For example Developers, Product Owners, Support, End Users ...'. Thus, testing is no longer the reserve of specialist testers and developers; instead many people are doing some testing as part of their role.

Although the IT industry has attempted to move towards the reduction of testing headcount by attempts to automate testing or by merging testing activities into other roles, reports such as MABL (2021); Practitest (2022); Capgemini (2022) show specialist testing skills are still required, as well as domain, business and UX testing skills. Indeed, the WQR from Capgemini (2022), whose respondents are selected from senior managers to C-level executives, remarks 'Testing efficiency is always top of people's minds [sic]' and also notes that it is 'both interesting and striking ... that the most important factor [for productivity] was deemed to be having enough staff with the right skills.' Capgemini (2022) indicates that the aspirations of C-level managers include specialist quality and testing roles for champions and experts on quality, customer experience and productivity. While Waychal et al. (2021) showed that practitioners who are not already testers may view testing as an undesirable career, perceiving the work as easy but tedious, they also found that people already engaged in software testing careers talked about the challenging, interesting nature of the work. The practitioners in Practitest (2022) and Capgemini (2022) view testing as a challenging job: 'Either unfortunately or fortunately, depending how you see it, a tester was and continues to be the equivalent of a Swiss army knife for his or her team. This means that we need to have a combination of skills that will serve our needs depending on the challenge being thrown at us. Just as always, flexibility, critical thinking, and the ability to learn all the time are the most necessary tools any tester needs to have to succeed in our trade.' (Practitest 2022, page 22). Petre and van der Hoek (2016) explore what it means to be an expert in software design, and one of their markers of an expert is that 'experts test'. Of course, that does not mean that every tester is an expert, but it is an indicator of the importance of testing.

2.3. HCI and UX

The study of Human-Computer Interaction (HCI) and the practice of User Experience (UX) design and testing are closely related; both focus on a better understanding of people's contexts, characteristics, goals and needs (Norman and Nielsen (2019)). Applying those techniques to understanding software testers' characteristics and contexts could be both useful and instructive. Potential benefits include a better understanding of actual and potential diversity among testers, perhaps informing recruitment, onboarding, work practices and tool design.

HCI and UX use people-centred, qualitative methods to understand both the context and goals of those who will use a specific software tool, identifying their relevant characteristics that affect the software's design. One such method is persona development, a rigorous empirically-driven process enabling designers to reason about who will be using their product or system, and take design decisions. Dam and Siang (2022) say personas are 'fictional characters, [based on research to] represent the different user types that might use your service, product, site, or brand'. Although typically, personas are used for software design, such a technique could also support activities such as recruitment, onboarding and tool acquisition for testers.

2.4. Stereotyping, Diversity and Hobbies

HCI and UX practices increasingly emphasise the need to study and understand a diversity of potential and actual people who will be using the software (Himmelsbach et al. (2019)). In an article on testing AI, Shneiderman (2022) suggests that diversity in teams brings benefits: 'muddy boots pragmatists' looking for risks, problems and solutions to reduce harm and remove bias must be part of the design process. Their article concludes 'The diverse workers of this camp ... have important messages to ensure that the blue-sky dreams can be channeled into realizable products and services that benefit people and preserve the environment.' Capretz et al. (2015) note that given the wide range of backgrounds, personalities and preferences of people using software, a wider range of people may be needed for software testing. If the C-level in organisations needs to recruit testers with both the right attitudes and the potential to rise to the organisation's aspirations for quality, then recruiting a widely diverse, interesting group of people to these challenging roles is important to the success of software development and delivery.

van Tilburg et al. (2022) discuss attitudes toward people perceived as boring and note that boring people are perceived as both emotionally cold and often as less competent. They note that certain jobs and hobbies are seen as boring, and therefore, the people who do those jobs are assumed to be boring. This is noteworthy for us; we examined hobbies as a way to understand the testers in our sample.

McChesney et al. (2022), the authors examine stereotypes as expressed in the O*NET, a mechanism for assessing role suitability by reference to hobbies and interests. O*NET classifies people according to set categories of interests: Realistic, Investigative, Artistic, Social, Enterprising and Conventional (RIASEC). Their study sampled 1000 actual and aspirant computer science workers and used the RIASEC scale to assess their interests. In O*NET the typical computer science worker is expected to score low for artistic and social activities. However, McChesney et al. (2022) suggest that this stereotype blocks entry from more diverse groups.

3. METHODOLOGY

This paper covers work to study the industry practitioners who are testing, their backgrounds and characteristics, based on the research question: **'Who is testing?'**, with data collected via a survey.

3.1. Survey Design and Piloting

We designed a detailed anonymous online questionnaire with a mix of open and multiple-choice questions. Our earlier studies Evans et al. (2020, 2021) with over 100 testers had identified emerging themes around tester backgrounds and roles, which enabled development of our survey. We also used literature such as McChesney et al. (2022) to identify relevant quotes, codes and themes to help with question design and data analysis. The survey had three main sections: participant backgrounds, their roles, and their activities. Survey questions are in the replication pack available on OSF[2] Background questions covered demographics, plus open-ended questions about backgrounds, hobbies, qualifications and training. The section on Roles covered years of experience in IT, job description, additional responsibilities, and how closely their job title matched their actual work. The Activities section asked open-ended questions about the approaches, techniques and tools they used in their work.

We reviewed the survey within our research team and piloted it with four testers recruited from an online tester forum, by asking for volunteers. The four pilot participants came from different countries, both native English speakers and with English as a second language. The survey was refined from their comments on and responses to the survey.

3.2. Participant Recruitment

We issued an open invitation via industry conferences and on-line technical discussion groups. This provided convenience as well as snowball sampling with a focus on practitioners, and the opt-in recruitment especially via conferences, Slack communities, webinars and meet-ups meant that the participants were not chosen by the researchers but self-recruited from the industry community. The population made up from these groups are international, and the requests to participate were on-line with a potential global reach, and in-person at events in Europe, North America and Australasia. Industry practitioners then shared the link to the survey with colleagues. Participation was welcomed from those who had a role with a large focus on testing and from those with testing as a secondary part of their role. Many participants identified as QA analysts, Software Engineers in Test, Test Managers, Developers, Project Managers, and Product Owners.

All participants were asked for consent and given the option to leave the study, either with the data so far available for use or asking for their data to be removed completely. They were also given the chance to say if they wanted to be further involved, and if so, to leave contact details. The data was then

cleaned, anonymising all participants and removing those who had withdrawn consent part-way through.

As a validity check on the sample, to understand the global population of testers and guard against bias, a comparison was made with industry surveys, including the World Quality Report from Capgemini (2022), media packs from two major conferences, Techwell (2022); EuroSTAR (2022), and two vendor industry surveys, MABL (2021); Practitest (2022). We found similarities in demographics to our sample.

3.3. Analysis and Coding

As well as examining closed questions, for example the demographics, using quantitative methods such as frequency analysis, we used qualitative analysis of open questions, looking for themes in the testers' responses. We adopted an interpretivist stance after Baltes and Ralph (2022), based on interpretation of qualitative data from open questions.

The analysis and coding looked for codes and, in turn, groups thereof in the answers to open questions. These included questions on hobbies, qualifications, training, job titles, roles and the domain in which the participant works. Example codes are in Table 1. For academic qualifications, answers were coded using terms such as 'masters', 'Arts', 'Science', and so on. Similarly, for training in testing, we adopted generic codes such as 'named courses' to code formal testing-related training such as ISTQB and BBST, and 'named experts', when participants mentioned experts they follow. Role descriptions were coded in terms of responsibilities (e.g., 'automation design', 'risk assessment'), and we also looked at people's writing styles (e.g. 'terse' versus 'expansive') and their approaches to testing and techniques (e.g., 'exploratory', 'state transition', 'happy path'). Furthermore, for answers around hobbies, we based our coding on McChesney et al. (2022); van Tilburg et al. (2022), using codes such as 'Artistic', 'Multi-Interested' and so on. These provided a rationale for the classification of hobbies, and an opportunity to re-code our data and validate our codes. Similarly, Capgemini (2022) and Practitest (2022) informed domain and sector codes.

Additionally, answers to open questions in the survey provided the participants' own words, often with rich stories of testers' work and home lives. We extracted interesting quotes from these responses, generating codes which built into themes such as 'enthusiasm', 'aspirations', and 'meeting challenges' which we found emerging from the stories. We categorised the responses from both open and closed questions into groups such as 'type of training' and 'academic qualification' to provide frequency counts.

Initial analysis was performed by one of the researchers, with data and emerging findings cross-validated by the other authors, in peer-reviews at industry conferences and examination of industry reports, standards and literature.

4. RESULTS

We summarise the results of our analysis in terms of both the demographic data and the coding of open questions, concentrating on who is testing, and the themes emerging from their backgrounds, home lives and work. We show evidence from our sample that indicates an interesting and diverse group of people, doing what they perceive as challenging work. There is potential for developing a diverse persona set, for example, to support test tool design.

4.1. Understanding and Validating the Sample

Age, gender and location were primarily collected to check for the sample having a reasonable match to the demographics of the industry, and, indeed, the general population[3]. Participants were of working age, with 53 per cent in the 25-45 age range. About 48 per cent of all respondents were male and just over 50 per cent female, with just over 1 per cent non-binary. The work of Waychal et al. (2021) suggests that studies of testing practitioners looking at location may provide further insights into their motivations and attitudes to testing. In our sample, 61 per cent of the participants were from Europe, 15 per cent from North America, 15 per cent from Asia, 4 per cent from South America and 3 per cent from Australasia. Although various efforts were made to reach out to local test communities in Africa, we did not get any responses from that continent. The industry domains that our participants worked in were checked against the sectors and domains covered by Capgemini (2022) and Practitest (2022). There was a matching diversity of industries in our sample, including technology, financial services, healthcare, and government.

4.2. Academic Qualifications and Training

The majority (85 per cent) of the survey participants had either a bachelor degree, post graduate qualification, or an equivalent, such as a vocational qualification to degree level. Just over 70 per cent of the respondents had attended formal training courses for testing. Other training options commonly mentioned included independent learning (25 per cent), company-based training (7 per cent), community-based learning (5 per cent) and following experts. We return to a discussion of learning and aspirations in Section 5.

Table 1: *Survey Topics - Qualitative Analysis from Open Questions*

Question topic	How categorised and coded
Hobbies	Coded to categorise by topic, level of activity, location, sociability, and recoded against codes and themes used in McChesney et al. (2022), including Artistic; Social; Stereotypical IT worker (low social, low arts).
Academic qualification	Coded to 5 level groups - (Doctorate, Masters, Bachelor, Pre-bachelor, not given) and Subject groups (Arts, Social Science, Science, IT, Other).
Training in testing	Code groups emerged from data as: 'named courses' e.g. ISTQB, BBST, 'named experts' and 'self-taught'.
Role Description	Coded in 3 groups (1) 'responsibilities' e.g. automation design, risk assessment, (2) 'approaches to testing' e.g. exploratory, scripted, and (3) 'techniques' e.g. state transition, happy path.
Aspirations	Coded as Low e.g. no wish to change, Medium e.g. looking to upgrade skills, run team or High e.g. CTO, international speaker
All open questions	Communication style e.g. 'terse', 'conversational', 'expansive', 'systematic'.

In Table 2 we show the breakdown of degree subject area (where known) and roles, approaches and responses style. The table shows that the degree subject is not a predictor of the role that someone might take; nearly half of the Arts and Social Science graduates were in technical roles, such as automation engineer, while nearly half of IT graduates were in non-technical roles. We also found that the response styles aligned to some extent with background: the Arts graduates wrote in detail, conversationally, while Science graduates tended to provide ordered, methodical lists of tasks, and IT graduates provided detail, but in a terse style.

4.3. Roles and Years of Experience

Our participants were mainly working at practitioner, senior practitioner and management levels, with a small number of junior testers, and a small number of director/C-level participants. 53 per cent of the participants mentioned previous roles or backgrounds. Of those, nearly half had come from a background outside IT or computer science. Under 20 per cent had spent their whole career in testing. Previous roles that involve attention to detail and clear communication would seem to match aspects of the tester's analysis and communication roles, for example, 'Prior to becoming a software tester I was a science communicator - my daily work highlighted the relevance of remotely sensed geophysical (atmospheric and terrestrial) data products using clear words and striking images' [SP46].

Table 3 summarises the years of experience in different groups of IT roles. As we might expect from the age ranges, the experience of testing and test management peaked between the 5-10 year range. People with longer experience had moved, as might be expected, into leadership and coaching roles.

In addition to years of experience, our open questions on approaches to testing which gave insight into changes with experience. For instance, Petre and van der Hoek (2016) show experts use visualisation; we found participants with longer experience use visualisation, whereas less experienced participants did not: 'I diagram or mind map just about everything' [SP1], and 'flow diagrams, context diagrams, mind maps, story mapping' [SP68].

4.4. Aspirations, Job Titles and Roles

Job titles often reflect actual roles; 66 per cent said their job title was reasonably accurate. However, even in that group, there were doubts: a Junior Test Engineer remarking 'the word engineer makes me wonder...' [SP3], while another participant notes 'Every test specialist in the company is called a QA Engineer, but with different actual focus . . . ' [SP14].

Many of the participants had multifaceted roles; for example, one mentioned 'I guess the only thing is that I have often had to wear a lot of hats' [SP16] while another went into detail: 'My expertise lies in test automation and testing. . . . lead of a team of 4 test automation specialists . . . on two occasions I have combined my testing work with a scrum master role . . . also have coached [testing and automation]. In my latest role, I am picking up ops tasks . . . My roles have required me to communicate a lot within my team, but have also involved speaking with the business, getting their feedback. I've also been heavily involved in guilds, communities, etc, setting up workshops. events and meet-ups.' [SP37].

The participants were generally full of aspirations for personal growth within testing as a career. Only 16 per cent expressed no aspiration to grow within the industry, including 2 per cent who are over 60 years old and talking about retirement, and 1 per cent who were simply happy in their current role: 'nothing. i love my current work and my team. I want to be better what im doing now' [sic] [SP27]. In contrast, many had aspirations to grow within their

Table 2: *Academic Subjects, Roles and Approaches*

Subject Area	% of Participants	Roles and Approaches
Arts	9.9%	42% in technical roles, such as automation engineer, and 28% in management roles. Conversational response styles. Multi-functional roles, team problem solving, diagramming.
Social Science	14.1%	40% in technical roles, 50% in management, 30% consultancy/coaching. Conversational response styles. Multi-functional roles, seeing the 'big picture' in a project.
Sciences	25.4%	44% in technical roles, 50% in management, 27% in non-technical roles. Terse response style, often as ordered lists.
IT	16.9%	58% technical role, 41% non-technical role, 50% had management responsibility. Response styles were to the point, not conversational, but full of technical detail.
Not stated or Pre-Bachelors	42.3%	31% in technical roles, 47% in management, 31% in non technical roles and 1.5% consultants. With an expansive, conversational style, this group had plenty of stories to tell about communication, pairing, and project-wide/non-testing aspects to their roles. This group used frameworks and methodologies when describing work approaches.

Table 3: *Experience in IT Roles*

IT Roles	$\leq$ 2 years	2-5 years	5-10 years	11-20 years	21+ years
Requirements, analysis and design	6%	24%	27%	10%	7%
Development and test automation	23%	25%	8%	6%	7%
Testing, test management	4%	14%	37%	24%	18%
Release, operations and support	20%	27%	13%	1%	6%
Leadership, risk, coaching, other	11%	23%	23%	15%	8%

current role, or could see a step up in their role; 66 per cent talked about upskilling, immediate team leading, and aspirations in terms of technical and business impact. For example, participant [SP67] is focused on the customer 'My goal is to ensure we deliver good quality - aspirations to learn the business and excel at it'. Participant [SP69] wants to 'Keep learning and adapting to the changing world in testing and maintain a position in Quality Assurance' and understands the need for lifelong learning simply to keep pace with technology. Participant [SP37] is clear that their life is about developing technically: 'I want to develop my C# skills, I want to integrate feminism more in my day to day work, I want to learn more about AI, no management aspirations . . . '. In this group, there seems to be no correlation between aspiration and either gender or age. Some participants (12 per cent) had aspirations to C-level roles and influential roles across the industry. These participants mentioned CEO (Chief Executive Officer), CTO (Chief Technology Officer) and CQO (Chief Quality Officer) roles.

Many participants expressed a desire to help others - locally in their teams, in their organisations, and their customers. Several also talked about aspirations to international conference speaking, or sought to influence across the industry; responses to these questions showed enthusiasm, aspirations, and a lack of boredom about testing.

4.5. Hobbies and Interests

Hobbies and interests were coded into a number of groups (see Table 4), and the frequency of each group counted. The most mentioned group was related to sport, exercise, and outdoor pursuits, including running, gardening and bird watching. The second most mentioned group was the creative group: makers, performers, writers and musicians.

Additionally, each response was examined to differentiate between hobbies that are (a) more passive and more active, (b) more indoor and more outdoor, (c) more arts or more STEM, and (d) more individual and more team-based (see Table 5). For each mention, a subjective score between 1 and 10 is given based on the content of their response (e.g. a score of 10 for the Indoor/Outdoor range is given if all the hobbies mentioned can only be carried out outdoors). Participants could be in more than one group, for example, with active artistic and passive STEM hobbies. An example of a passive indoor hobby is one participant who simply put 'watch netflix' [SP11], while an active, multi-faceted response was 'i love music and non-mainstream cinema; i travel a lot and go to music festival; i'm passionate about people and love to spend time with my friends; i'm very active, always doing stuff and planning new adventures' [sic] [SP3]. An example of a more STEM-related hobby was 'Photography, forensic anthropology

Table 4: *Hobby Groups: Percentage of Respondents*

Hobby Group	Survey %
Sport, exercise, fitness, outdoors	60%
Making, performing, writing, playing	45%
Reading, studying, learning	41%
Watching/listening theatre/music	28%
Family and friends	26%
Watching film, cinema and TV	17%
Games and puzzles	17%
Food and drink	15%
Travel, other cultures	13%
Volunteering and community	8%

and medical books, word games, escape rooms, travel' [SP15] while 'Painting, kickboxing, feminism, reading, dancing, volunteering, making costumes, creating in general' [SP37] is someone active in arts and community. The group scored more highly for active, indoor, and artistic hobbies and had a balance between individual and team hobbies. We cover this more in Section 5, where we compare our results with those in McChesney et al. (2022).

5. DISCUSSION

This work was driven by the question: '**Who is Testing?**.' We found a wide diversity of people and backgrounds in industry. The emerging picture is complex: it is an interwoven tapestry of variable characteristics, affecting how a person approaches testing, their motivations and the support they need.

5.1. Who is Testing?

5.1.1. Diversity of Backgrounds

The diversity of backgrounds and experiences that testers bring to teams is evidenced in the data we have collected. Although a number of respondents had entered testing by what might be seen as a conventional route of a software engineering degree, testers also entered their roles with other degrees, both science and arts, or no degree at all (see Table 2). Unexpected backgrounds included philosophy, medicine, theatre, music, physics, HR, urban planning, pharmaceuticals and boat building: ' . . . realised I didn't want to be a boat builder after all so . . . took up testing . . . that's where I've been ever since . . . ' [SP42].

5.1.2. Diversity of Hobbies

This diversity continued when we examined testers' hobbies in the survey, with testers listing a wide range of types of interests, with many actively engaged in making and doing. For example, many respondents listen to music (29 per cent mentioned it specifically), and some are making music (12 per cent singing or playing), and even in one case composing. Creative visual design and making were frequently mentioned (27 per cent), with outdoor and exercise-related hobbies the most frequently mentioned: gardening to horse riding, running and kickboxing (61 per cent).

We compared McChesney et al. (2022)'s figures with the hobbies information we collected from software testers in the survey (see Table 6), and re-coded against their themes, finding testers to score highly on practical (66 per cent) and arts (91 per cent), with people generally having a range of interests (74 per cent). This was not connected to degree subject; the pattern was similar across subjects. Software testers appear to be more artistic and to have a wider range of interests than the group of IT workers McChesney et al. (2022) studied, who themselves do not conform to the O*NET stereotype for IT workers. Only 6 per cent of our participants conformed to the O*NET stereotype for IT professionals, and the range of hobbies broke outside the boring' stereotypes in van Tilburg et al. (2022). There did not appear to be gender bias; roughly half of the artistic participants in our work identified as male. McChesney et al. (2022) do not report on all the ranges of activities covered in our work, shown in Table 6.

Hobbies such as singing in a choir or playing in a sports team are indicative of an orientation to team based work activities such as ensemble or pair testing. The O*NET stereotype of an IT worker's interests was for a low score on social activities, indicating a solitary nature. The range of communication styles and hobbies (Tables 2, 5 and 6) show a group of people together able to meet the challenges of teamwork, remote working and customer focus within modern organisations.

We saw a pattern of divergence from the perceived stereotypical IT worker described in both O*NET and McChesney et al. (2022) work; our tester participants reported many social engagements, with family, friends, fellow hobbyists and in the testing community. Only 6 per cent of our participants conformed to the O*NET stereotype, compared with 30 per cent of McChesney et al. (2022)'s sample: we expect most of our participants' work behaviours and communication style to be non-stereotypical.

5.1.3. Learning and Aspirations

We also found in our data that testers have the ambition to seek improvements in their own work, and also in the software and systems they help deliver. For example, participants' responses about testing training could be grouped into several areas: courses based on named syllabi or schemes, courses presented by particular experts, community-based, and self-taught/on-the-job training. For example, one mentioned 'Effective Methods of System Testing, ... Rapid Software Testing; Getting

Table 5: *Range of Hobbies Mentioned - Heatmap**

Min	1	2	3	4	5	6	7	8	9	10	**Max**
Passive	1	2	2	4	13	19	15	13	3	0	**Active**
Indoor	2	10	6	10	7	22	10	3	0	0	**Outdoor**
Art	2	4	20	27	13	2	1	2	0	0	**STEM**
Individual	4	14	13	7	15	6	10	2	0	0	**Team**

* Darker color indicates greater number of participants scoring on scale point.

Table 6: *Tester Hobby Types Not Stereotypical (after McChesney et al. (2022))*

	McChesney	McChesney	Our survey
Sample Size:	500	500	71
CS/IT role:	Employed	Aspiring to	SWT
Artistic	20% (72%*)	41% (60%*)	91% (53%*)
Practical	not given	not given	66%(51%*)
Analytic	not given	not given	38% (48%*)
Social	not given	not given	28% (32%*)
Multi-interested	19% (27%*)	not given	74% (54%*)
Few interests	31% (31%*)	23% (44%*)	23% (35%*)
Stereotypical IT	30% (30%*)	36% (36%*)	6% (50%*)

* Of which are female/non-binary (after McChesney et al. (2022)).

a Grip on Exploratory Testing; Black Box Software Testing (BBST) Foundations' [SP1], while another listed 'ISTQB Certified Tester Foundation Level, BCS Requirements Engineering' [SP58].

Micallef et al. (2016) discusses how training impacts testers; we found over 50 per cent of those who discussed training appreciated a mix of ways of learning, for example: 'Maaret on Exploratory Testing Foundations course . . . one of the best trainings . . . loved Fiona Charles's masterclass on Test Strategy. . . . hours and hours of self-learning . . . attending conferences . . . practicing.' [SP2].

This emphasis on learning may be of significance considering the call from C-level executives in Capgemini (2022) for greater skill levels among testers as the **main factor in improving productivity.** Learning also reflects the testers' stated aspirations reaching to an ambition for C-level roles, and extend to wanting industry-wide influence, stepping up to the challenges laid down in Capgemini (2022) and Practitest (2022).

5.2. Are Testers Different?

Our data suggests that testers possess a diverse range of interests and display a distinct profile. In contrast to participants in McChesney et al. (2022), where 30 per cent met the O*NET stereotype, only 6 per cent of our testers were stereotypical: we found testers tend to have multiple interests, engage in artistic and outgoing hobbies, deviating significantly from the stereotypical classification found in O*NET and in McChesney et al. (2022).

We also found indications in our own work (Table 2) and literature (Hermans (2021)) that variation of background may bring different communication styles and approaches to work. This informs organisational culture, teamwork and cooperation, encouraging better information exchange. Whereas software developers tend to originate from technical training, our data shows that testers' education is rooted in a variety of different fields. The barriers to entry for testing seem to be lower than for other roles. For example, whereas Cavin (2015) reports on successful efforts training veterans to take up roles in software testing, a similar initiative for doctors in the UK to learn coding in Hoeksma (2013) seems less successful.

We argue that the ability of the profession to attract people with such diverse backgrounds stems from its multi-faceted and distinct nature. Our data shows a spectrum of roles and responsibilities within the field, while Hernández and Marsden (2014) found that 91 per cent of its participants were attracted to the job due to its variety. A study amongst 5,971 developers by Meyer et al. (2019) showed that most of a developer's day is spent coding (17 per cent), bug fixing (14 per cent), attending meetings (15 per cent) and answering emails (10 per cent). On the contrary, Stray et al. (2022) observe that testers spend half their day communicating and learning how to mitigate the fact that they have to bring bad news to other project members. They also invest considerable effort in facilitating communication between different types of stakeholders. This is in addition to designing tests, implementing test frameworks, executing tests and managing the test process. This is echoed in our

participants' responses, for example, 'My roles have required me to communicate a lot within my team, but have also involved speaking with the business, getting their feedback.' [SP37].

Therefore whilst software developers are likely to have similarly diverse interests to software testers, the latter cohort is more diverse in terms of its members' origins and the range of tasks that they need to complete on a day-to-day basis.

5.3. Why This Matters

We now reflect on why our results matter for society, for recruitment and on-boarding of testers, and for work practices and tools used in testing. We draw on the concepts discussed in Section 2 as well as our results. We indicate that stereotyping might be problematic, and highlight the importance of diversity. We note that testing is challenging, requiring a diverse group of testers and that the design of work practices and tools could be supported by HCI approaches such as personas.

5.3.1. Society

We noted in Section 2 that Himmelsbach et al. (2019) and Shneiderman (2022) represent HCI and AI communities increasingly calling for consideration of diversity. Supporting the testing community as a diverse and heterogeneous group helps encourage the notion of diversity in the IT industry. The warning from Himmelsbach et al. (2019) that 'applying categories can bear the risk of labelling users as 'the other' ' means care is needed to ensure that a diverse group of testers are neither ignored, nor encouraged to become a homogeneous and more stereotypical group of IT workers.

Product designers need to understand and support a heterogeneous and diverse customer population, and that is supported by a heterogeneous tester population representing them (Himmelsbach et al. (2019); McChesney et al. (2022)). Our results show that a diversity of background exists in software testing, with a much closer match to the general population than studies of other IT practitioners. However, it is possible increase the diversity further, and fully reflect the general population.

However, it is important to distinguish stereotyping from acknowledging diversity, whether looking at roles, cultures, gender, or other factors.

5.3.2. Recruitment, On-boarding and Training

We have shown in this paper that the recruitment of testers is potentially problematic, because of stereotyping and perceptions of tedium (McChesney et al. (2022); van Tilburg et al. (2022); Waychal et al. (2021)). We have also shown that testers come from diverse backgrounds, and that this rich mix of personas, together with the challenging and multifaceted roles they perform, could incentivise a wider population to take up a career in testing, to the benefit of society. Shneiderman (2022); Capgemini (2022); Practitest (2022) all state that a diverse and multi-skilled group is essential. Providing an understanding that the group already in the industry is diverse may engage and encourage even more diverse groups to apply, and to be recruited.

The skills, training and development required by testers throughout their careers follows from their recruitment to staff retention and development. Studies by Capretz et al. (2015); Weilemann (2019) have identified personality traits that best fit the requirements of various roles and tasks in software engineering, including testing. Our data indicates that such efforts may be misleading.

Hermans (2021) discusses the use of numeracy and language abilities as predictors for programming ability, with the comment: 'This is interesting because we as a field typically stress the fact that mathematical skills are important' and the finding that **language abilities were actually the better predictor of success at programming**. A wider ranging recruitment policy both in colleges and by employers could benefit both the industry and society, but increasing both programming, testing and communication skills.

Capgemini (2022) says C-level executives have aspirations for the test community as 'custodians of . . . quality at speed . . . focusing on value and on avoiding the possibility of defects'; testers' diversity of thought and productivity to support this could be fueled by methods and tools that support efficiency, effectiveness and preferred ways of working.

5.3.3. Test Work Practices and Tools

The maturity of UX practices described by Nielsen (2006) is partly judged by the understanding of the people for whom a tool is designed. With a heterogeneous target group, this can provide challenges to the designer. Wöckl et al. (2012) discusses the use of a basic set of personas to cover widely diverse groups in what can, to outsiders, look like a homogeneous group. In their case, the group is older adults, and in our case, it is software testers. The development and use of a set of personas allows an initial understanding of the target group. For testers, the personas could include themes and groupings we identified in this study, such as communication styles and teamwork preferences (see Sections 5.1.2 and 4.2).

The responses from our participants indicate that communication needs and styles among testers differ more widely than those of other IT workers

(see Tables 4, 5, and 6). The mix of practical, analytic and artistic hobbies indicates potentially a wide range of ways of communication styles and problem-solving preferences (see Table 6). Both tool interface design and workflow design would be affected by broadening understanding of tester personas to include preferences we found. For example, a team oriented, social person will favour workflows and tools that support pair and group problem solving; not all test tools support this well (Evans et al. 2020). Petre and van der Hoek (2016) show tool interfaces with strong visual elements may suit experts; we add, especially those with a leaning toward visual arts. A spreadsheet-style interface may be more suitable for testers with an analytic preference, indicated by hobbies.

The characteristics of testers, their backgrounds, capabilities and needs inform the characteristics of the workflows and tools they require. Automation won't remove the need for testing skills. In fact, all the indicators are that testers with increased skills will be given greater and more wide-reaching responsibilities for overall quality. Current tool support for testing presents testers with challenges Evans et al. (2020). Testing experts recognise this; For example, MABL (2021) remarks: 'We all recognize that . . . quality professionals play an essential role in building market-leading products and services. . . . A diverse group of people will be needed to carry out those challenging activities' (MABL (2021); Practitest (2022)). Their activities must be supported by workflows and a toolset that enables them to carry out their activities with effectiveness and efficiency, maximising their productivity and capability to become the champions of quality while meeting their own C-level aspirations.

The need to support for a variety of people with different characteristics is supported by our own work, and also by industry reports; 'It's clear that QA will lead testing, but it's also beneficial to work toward democratizing testing with an everyone involved approach to create a culture of quality.' (MABL (2021)). Research from Waychal et al. (2021), McChesney et al. (2022) and van Tilburg et al. (2022) indicate why the vital role of testing may not be popular, while this paper may counter those perceptions positively.

6. CONCLUSIONS

The activity of software testing is challenging, requiring continuous learning, problem-solving, critical thinking and flexibility. It is not narrowly focused on technical testing, as the role of testers is changing, with a widening of responsibility and influence.

Who is testing? Testers are a diverse, interesting, and lively group of people with aspirations to improve their organisations and themselves, to better support their customers. Using qualitative data collected from industry practitioners, we have shown that testers are not stereotypical IT workers, but are heterogeneous. This is a benefit that meets their technical, organisational, and societal challenges. Whether artists, engineers, or computer scientists, they do not conform to stereotype, and require support, workflows and tools that enhance their unique challenges, preferences and capabilities.

We have shown that testers are well placed to meet the expectations of C-level executives quoted by Capgemini (2022) in terms of influence and argumentation for increased quality at all levels in their organisation. The diversity we demonstrate helps meet the ethical challenges of technologies, such as AI, and supports Shneiderman (2022). The range of skills required indicates teams made of heterogeneous group of individuals. Encouraging entry to the industry from as wide a group as possible is healthy for the testing industry, the wider IT industry, and society. Breaking the stereotype strengthens the industry's ability to deliver what is needed by society and by customers.

Our conclusions do not claim to establish a pattern in the population. Quite the contrary, our results indicate that the testing community is a diverse one which cannot be pigeonholed into one bucket called 'testers'. With regards to internal validity, we consider the methodology discussed in Section 3 independently replicable and believe that the results discussed represent the characteristics of our specific cohort of participants. This claim follows from the research protocol which included systematic efforts of data cleaning, coding and most importantly, regular validation with the community.

We identify several areas of interest for future research. Understanding the range of test personas, their work styles and communication preferences, informs IT methodologies and tool design. Further understanding of stereotyping would help organisations reflect the general population when recruiting testers, positively impacting recruitment and retention from a wider range of backgrounds including academic background, work experiences and hobbies. This would fit the increasingly high profile, complex and ethically essential testing roles required by industry and society.

REFERENCES

Bach, J. and Bolton, M. (2016). A context driven approach to automation in testing.
URL: *https://www.satisfice.com/download/a-context-driven-approach-to-automation-in-testing*

Baltes, S. and Ralph, P. (2022). Sampling in software engineering research: A critical review and guidelines, *Empirical Software Engineering* **27**(4): 94.

Beizer, B. (1984). *Software system testing and quality assurance*, Van Nostrand Reinhold Co., New York.

Capgemini (2022). World Quality Report 2021-22.
URL: *https://www.sogeti.com/explore/reports/world-quality-report-2021-22/*

Capretz, L. F., Varona, D. and Raza, A. (2015). Influence of personality types in software tasks choices, *Computers in Human behavior* **52**: 373–378.

Cavin, J. D. (2015). The role of human factors in veteran sqa training, *Procedia Manufacturing* **3**: 1535–1542.

Dam, R. F. and Siang, T. Y. (2022). Personas: A simple introduction.
URL: *https://www.interaction-design.org/literature/article/personas-why-and-how-you-should-use-them*

EuroSTAR (2022). Eurostar Media Kit 2022.
URL: *https://conference.eurostarsoftwaretesting.com/expo/*

Evans, I., Porter, C. and Micallef, M. (2021). Scared, frustrated and quietly proud: Testers' lived experience of tools and automation, *European Conference on Cognitive Ergonomics 2021*, ACM, New York, NY, USA, pp. 1–7.

Evans, I., Porter, C., Micallef, M. and Harty, J. (2020). Test tools: an illusion of usability?, *2020 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, IEEE, IEEE Computer Society, Washington, DC, USA, pp. 392–397.

Hermans, F. (2021). *The Programmer's Brain: What Every Programmer Needs to Know about Cognition*, Simon and Schuster.

Hernández, T. P. R. and Marsden, N. (2014). Understanding software testers in the automotive industry a mixed-method case study, *2014 9th International Conference on Software Engineering and Applications (ICSOFT-EA)*, IEEE, pp. 305–314.

Himmelsbach, J., Schwarz, S., Gerdenitsch, C., Wais-Zechmann, B., Bobeth, J. and Tscheligi, M. (2019). Do we care about diversity in human computer interaction: a comprehensive content analysis on diversity dimensions in research, *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, New York, NY, USA, pp. 1–16.

Hoeksma, J. (2013). Turning doctors into coders, *Bmj* **347**.

ISO (2013). ISO/IEC/IEEE 29119-1: 2013 Software and systems engineering - Software testing - part 1: Concepts and definitions, *Standard*, International Organization for Standardization, Geneva, CH.

ISTQB (2023). International software testing qualifications board.
URL: *https://www.istqb.org/*

MABL (2021). State of Testing in DevOps 2021.
URL: *https://www.mabl.com/resources/2021-testing-in-devops-survey-report*

McChesney, J. E., Behrend, T. S. and Glosenberg, A. (2022). Stereotypical descriptions of computer science career interests are not representative of many computer scientists, *Scientific reports* **12**(1): 1–9.

Meyer, A. N., Barr, E. T., Bird, C. and Zimmermann, T. (2019). Today was a good day: The daily life of software developers, *IEEE Transactions on Software Engineering* **47**(5): 863–880.

Micallef, M., Porter, C. and Borg, A. (2016). Do exploratory testers need formal training? an investigation using hci techniques, *2016 IEEE Ninth International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pp. 305–314.

Nielsen, J. (2006). Corporate UX Maturity: Stages 1-4.
URL: *https://www.nngroup.com/articles/ux-maturity-stages-1-4*

Norman, D. and Nielsen, J. (2019). The definition of user experience (ux).
URL: *https://www.nngroup.com/articles/definition-user-experience/*

Petre, M. and van der Hoek, A. (2016). *Software Design Decoded: 66 Ways Experts Think*, MIT Press, Cambridge, MA.

Practitest (2022). State of testing report.
URL: *https://www.practitest.com/state-of-testing/*

Robert D. Austin and Gary P. Pisano (2017). Neurodiversity as a competitive advantage: Why you should embrace it in your workforce.
URL: *https://hbr.org/2017/05/neurodiversity-as-a-competitive-advantage*

Rodríguez-Pérez, G., Nadri, R. and Nagappan, M. (2021). Perceived diversity in software engineering: a systematic literature review, *Empirical Software Engineering* **26**: 1–38.

Shneiderman, B. (2022). Artificial intelligence needs both pragmatists and blue-sky visionaries.
URL: *https://www.scientificamerican.com/article/artificial-intelligence-needs-both-pragmatists-and-blue-sky-visionaries/*

Stray, V., Florea, R. and Paruch, L. (2022). Exploring human factors of the agile software tester, *Software Quality Journal* **30**(2): 455–481.

Techwell (2022). Techwell Media Kit 2022.
URL: *https://www.techwell.com/advertise-with-techwell/media-kit-form*

Tulonen, A. (2013). Project Mercury and the Death of Testing.
URL: *https://flowoftesting.wordpress.com/2013/12/21/project-mercury-and-death-of-testing*

van Tilburg, W. A. P., Igou, E. R. and Panjwani, M. (2022). Boring people: Stereotype characteristics, interpersonal attributions, and social reactions, *Personality & social psychology bulletin* **25**: 1461672221079104–1461672221079104.

Waychal, P., Capretz, L. F., Jia, J., Varona, D. and Lizama, Y. (2021). Practitioners' testimonials about software testing, *2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, IEEE, IEEE Computer Society, Washington, DC, USA, pp. 582–589.

Weilemann, E. (2019). A winning team-what personality has to do with software engineering, *2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion)*, IEEE, pp. 252–253.

Wöckl, B., Yildizoglu, U., Buber, I., Aparicio Diaz, B., Kruijff, E. and Tscheligi, M. (2012). Basic Senior Personas: A Representative Design Tool Covering the Spectrum of European Older Adults, *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, Association for Computing Machinery, New York, NY, USA, p. 25–32.

NOTES

[1] In this paper, we do not study aspects of diversity such as race, ethnicity and disability (Rodríguez-Pérez et al. (2021)).

[2] `https://osf.io/e6zjn/?view_only=13a99944d3534ed19dc5136ea48792b3`. Full datasets are not available as it is not possible to fully anonymise them (Baltes and Ralph (2022)).

[3] Based on data from `https://www.statista.com` (accessed 26 Sept 2023)