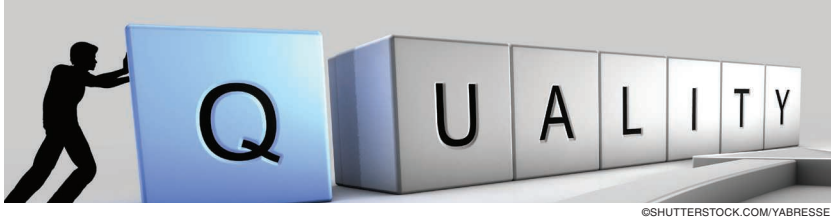
Test Maturity Model Integration: Trends of Worldwide Test Maturity and Certifications

Vahid Garousi, Queen's University Belfast and Bahar Software Engineering Consulting

Erik van Veenendaal, TMMi Foundation

// Test maturity model integration (TMMi) is popular for the maturity assessment and capability improvement of software testing practices. We present a status report for TMMi, including motivations and benefits, and discuss how companies have been ranked in its process areas. //



©SHUTTERSTOCK.COM/YABRESSE

IN RESPONSE TO the growing demand for software quality, various initiatives, models, and approaches have been presented in the software industry since the 1980s. Examples include the capability maturity model integration (CMMI) model (https://cmmiinstitute.com) and the ISO/International Electrotechnical Commission 15504 standard, also known as the *software process improvement and capability determination* (*SPICE*) *model*.

Over the past three decades, CMMI adoption has gradually increased, mostly for organizations working in governmental and defense projects. According to the CMMI Institute (https://cmmiinstitute.com/pars/) as of this writing (December 2020), 9,650 companies have received appraisals (certifications) for CMMI.

While studies have reported that models, such as CMMI and SPICE, could be useful in certain contexts, their primary focus is process improvements in the "overall" software development process. Such process improvement models do not often provide specific improvement recommendations for software testing. For this reason, various models have been developed for dedicated improvements of software testing practices. A 2018 survey article in *IEEE Software*[1] reported a catalog of 58 models for test maturity and capability improvements, e.g., test maturity model integration (TMMi) (https://www.tmmi.org).

In this article, the authors, one of whom is from the TMMi Foundation, provide a brief status report about the trends of worldwide TMMi assessments and certifications. Let us note that the TMMi model has been in existence since 2012, and, thus, the article's goal is certainly not to present TMMi but to discuss how companies have been ranked in each of the TMMi process areas (PAs) as well as the motivations for and benefits of using TMMi. The data have been compiled and prepared in an anonymous manner from the internal assessments and certifications database of the TMMi Foundation. We start by presenting a brief overview of TMMi.

Digital Object Identifier 10.1109/MS.2021.3061930
Date of current version: 14 February 2022

A Brief Overview of TMMi

The roots of TMMi goes back to Gelperin and Hetzel's evolutionary testing model,[2] published in 1988, and an early test improvement model named *TMM*.[3] Seeing the need for a more focused test improvement model, several test and quality experts (volunteers) came together (mainly based in Europe) in 2005 and formed the TMMi Foundation. The first version of the TMMi specification (1.0)[4] was presented by the Foundation in 2012. The latest version of specification, as of this writing, is 1.2[5] (released in 2018).

The TMMi framework[5] is a guideline and reference framework for test process improvement (TPI). TMMi uses the concept of maturity levels for process evaluation and improvement. Furthermore, for each maturity level, a set of PAs, goals, and practices is identified. The Foundation is supported by the so-called TMMi local chapters, which provide and organize TMMi-based consulting services locally in their country/region. At the time of this writing, 22 TMMi local chapters are in existence, e.g., in China, the United States, Spain, Brazil, and France.

TMMi is aligned with international testing standards and the syllabi and terminology of the International Software Testing Qualifications Board (ISTQB). The TMMi Foundation has consciously not introduced new or its own terminology but reuses the ISTQB terminology. This is an advantage for all test professionals who are ISTQB certified (approximately 700,000 worldwide at the time of this writing). With TMMi, organizations can have their test processes objectively evaluated by accredited assessors. It is also possible for test professionals and consultants to be personally certified as a "TMMi professional."

Similar to CMMI, TMMi has a "staged" scheme for test process assessment and improvement. It contains stages or levels through which an organization passes as its testing process evolves from one that is ad hoc, also called *initial* or *unmanaged* (level one), to one that is managed (level two), defined (level three), measured (level four), and optimized (level five).

Figure 1 shows the five maturity levels of TMMi, their 16 PAs, and the structure of TMMi as a metamodel. Achievement of any level by a given test team/organization requires that all PAs of that level and the lower levels have been satisfied.

In a hierarchical setting, each PA has several specific goals (SGs), specific practices (SPs), subpractices, generic goals (GGs), and generic practices (GPs). Across the five levels, there are, in total, 50 SGs, 173 SPs, 845 subpractices, 32 GGs, and 192 GPs. Details of those elements can be found in the TMMi specification.[5]

For instance, maturity level two ("managed") has five PAs, e.g., PA 2.1 (test policy and strategy). This PA has three SGs: SG 1 (establish a test policy), SG 2 (establish a test strategy), and SG 3 (establish test performance indicators). SG 1, in turn, has three SPs, e.g., SP 1.1 (define test goals) and SP 1.2 (define test policy).

A main underlying principle of TMMi is that it is a "generic" model applicable to various lifecycle models. The model has been translated into several languages, e.g., Spanish, French, Portuguese, and Chinese. Several experience reports and case studies from industrial applications of TMMi have been published, e.g., Garousi et al.,[1] van Veenendaal et al.,[6] Rungi and Matulevičius,[7] and Alone and Glocksien.[8]

For instance, one publication[7] reported a single-object case study of TPI using TMMi by an Estonian gaming software company. The study empirically found that the comprehensive and detailed documentation of the TMMi framework in terms of the reference model, assessment method, and data submission requirements provided adequate support in assessing and improving an organization.

In another study,[8] both TMMi and TPI-Next were applied in the context of a large Swedish company (Volvo IT). The study found that, even though the two models generally show strong similarities, differences in the assessment results are noticeable due to their different model representations. Mapping and comparison of the assessment results indicated that the requirements of the maturity levels in TMMi are much stricter and more difficult to reach than in TPI-Next.

Also, "the work examples in TMMi give very detailed descriptions of the testing process, which provides a good guidance in conducting the assessment."[8] Furthermore, the industrial study found that, for the successful application of both approaches, extended knowledge in software testing is essential.

A logical question to explore is about the motivations of companies to assess and improve their processes using TMMi. Maturity models that address the entire software development lifecycle, e.g., CMMI, provide only some high-level assessment and improvement criteria for software testing and, thus, are of limited use for TPI.

By surveying both the academic and gray literature, a study[1] reported that, in general, the three main drivers for utilizing TPI models are the need for increasing software quality, necessity to decrease the cost of testing, and process and operational requirements. A recent survey performed by the TMMi Foundation in 2020 (results are not published yet) shows that the main reasons

(motivations) for the adoption of TMMi are to enhance product quality, reduce product risk, increase testing productivity (efficiency), benchmark against an internationally used model, and increase the prestige of testing teams.

As discussed earlier, there are more than 58 models for TPI.[1] We were interested in objectively comparing the

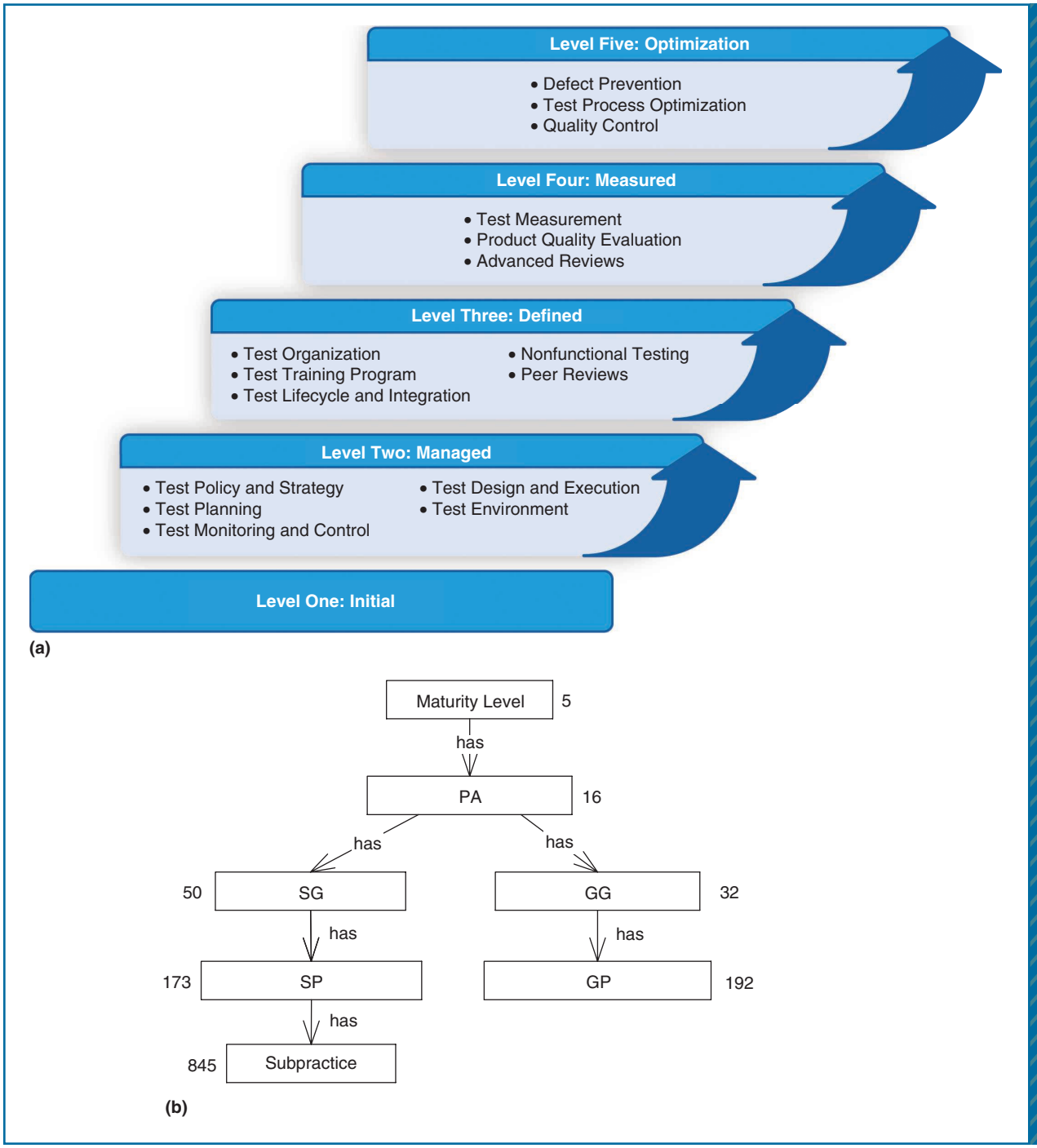


FIGURE 1. The (a) maturity levels of TMMi and their PAs and (b) structure of TMMi as a metamodel.

industry diffusion/penetration of different models. However, there are very limited data for this. In a survey study,[1] which reviewed sources from both industry and academia, 57 out of the 181 sources used TMMi, whereas 18 used a model named *TPI-Next*.[9] In the absence of objective statistics for popularity of maturity models in industry, these numbers could serve as popularity measures for TMMi and TPI-Next.

While TMMi is an independent model managed by the TMMi Foundation, the TPI-Next model has been developed by a software consulting company. We compare TMMi and TPI-Next in Table 1.

In terms of how TMMi compares with other existing models, an earlier study[1] discussed that TMMi and TPI-Next are both "general-purpose" models for TPIs, while there are "special-purpose" models, such as the unit test maturity model and automated software testing maturity model.[1] From the set of 58 models in this area,[1] it is not really easy to identify the most "promising" model from the point of view of the completeness of evaluations. Depending on its needs, a given team/company would choose and apply the right improvement model.

Table 1. A comparison of TMMi and TPI-Next.

Criteria	TMMi	TPI-Next
Representation	Staged model	Continuous model
Test levels	All test levels (unit, integration, system, user-acceptance testing)	Focus on higher test levels (system and user-acceptance testing)
Supported test methodology	Test-method independent	Linked to TMap (https://tmap.net)
Terminology	ISTQB based	TMap based
Base SPI model	Related to CMMI	None
Certification	Possible through formal assessment	None

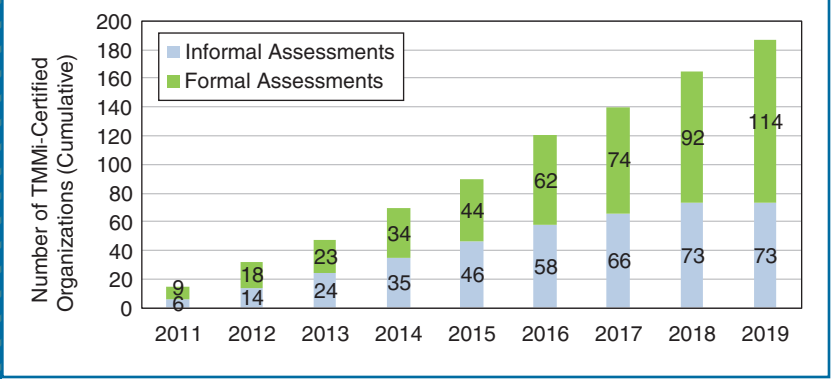


FIGURE 2. The growing number of TMMi assessments and certifications since 2011.

A Growing Number of Certifications

According to the internal certifications database of the TMMi Foundation, until the end of 2019, 187 companies worldwide submitted assessment applications to the Foundation. Of these, 73 companies applied for "informal" assessments, as their purpose was to get "an indicative rather than formal detailed maturity rating."[5] On the other hand, 114 companies applied for "formal" assessments since their need was to receive "a detailed assessment and/or [they] considered embarking on a path to certification." Out of those, 111 certificates have been issued. (The list can be found at https://www.tmmi.org.)

All assessments are systematically conducted by a team of certified TMMi (lead) assessors. For assessments, in addition to the specification,[5] there is another document named *TMMi Assessment Method Application Requirements* (*TAMAR*),[10] which is used to ensure rigor and consistency in assessments. Figure 2 illustrates the cumulative trend of TMMi assessments between 2011 and 2019, which shows an increasing interest in getting TMMi assessments.

When analyzing the assessments and certifications database, we were interested in gaining insight into the regions and countries that have had highest and lowest representation in TMMi assessments. In Figure 3, we show those data. The United Kingdom, South Korea, and China are the top three countries.

It is interesting that certain countries have had higher uptake rates. In the United Kingdom, for example, TMMi has been widely adopted in governmental units, and there are also media news articles[11] about it. This is partly due to the fact that the Foundation has made more efforts to

publicize the model in certain countries (due to the efforts of its members and local chapters). We should add that the Foundation started from the United Kingdom initially.

Furthermore, we feel that the work culture of certain countries (being "process oriented") and the large presence of CMMI in some regions are other factors impacting higher penetration of TMMi in certain countries. Note that some important IT-active countries are still missing in Figure 3. It is expected that this is soon to be changed. For example, three TMMi local chapters have recently been established in Canada, Italy, and Turkey. Others, such as Germany and Russia, are in the process of becoming a TMMi local chapter.

One way to make more sense of these country data is to correlate them with another metric in the Foundation's database, the locations (residence countries) of certified TMMi professionals. TMMi professional is a certification that is open to everyone with an interest in TMMi. Holding the TMMi professional certificate is a prerequisite to becoming an accredited TMMi assessor, the holders of which have the authority to conduct TMMi assessments.

Based on the compiled data, we show in Figure 4 the scatterplot of the number of TMMi professionals in countries (as of the end of 2019) versus the number of assessed organizations in those countries. Note that, for brevity, only countries with higher representation in terms of both metrics are shown in this figure (at least three in each metric). The Pearson correlation coefficient of the two series is 0.51, denoting a moderate correlation.

Thus, we can say that, generally speaking, the higher the number of TMMi professionals in a country, the greater the chance of having more assessed organizations in that country. This is indeed expected since, when there are more TMMi professionals in a country, they help increase the awareness for TMMi and also encourage more organizations to apply for assessments and certifications.

We also see other various interesting insights in the scatterplot of Figure 4; e.g., India has the third-largest TMMi professional community but ranks low in the number of TMMi-certified organizations. It could be that, in certain countries, companies do not see enough motivations/reasons to get certifications, e.g., if there is no enforcement by governmental agencies, unlike the case of United Kingdom discussed earlier. There may be many professionals doing TMMi-related work, but for some organizations, "just" achieving benefits, e.g., in terms of software quality, is often sufficient, and they may not go as far as getting the certification.

TMMi Levels Achieved by Companies

The other data in the certifications database are the "grades" for each PA and SG achieved by each applicant company or organizational unit. In Figure 5, we show an "individual-value" plot

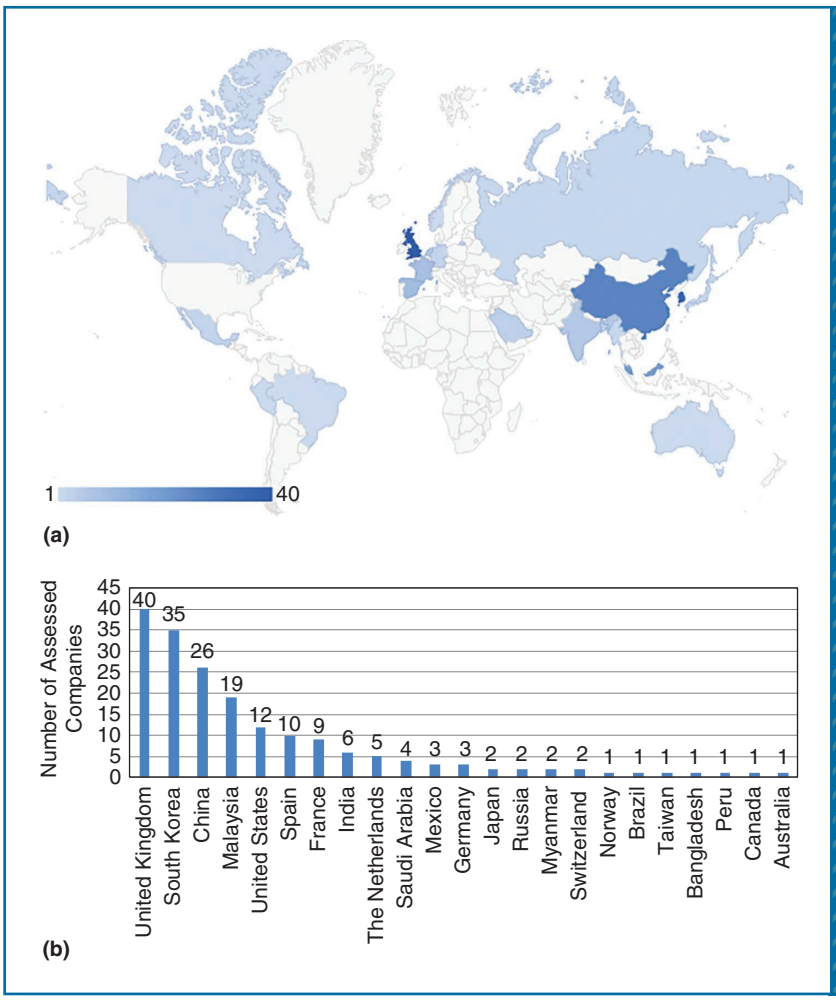


FIGURE 3. The (a) countries (locations) of the organizations that have received TMMi assessments and (b) number of assessed companies per country.

including the "moving" averages (length = 10 index of data points) of the TMMi levels achieved in all certifications (N = 111). We have also looked at the median and mean values for each year.

As we can see, the moving averages mostly fluctuated between three and four in the window of eight years. For example, during the year 2011, eight certificates were issued to companies with rather low TMMi levels (between one and three).

Overall, we do not observe any increasing trend in TMMi levels achieved in the certifications over the years. The main reason is that, in a given year, any number of companies with any level of maturity could have applied and received certifications in various levels. Also, the data set of maturity levels is not for the same set of companies over different years.

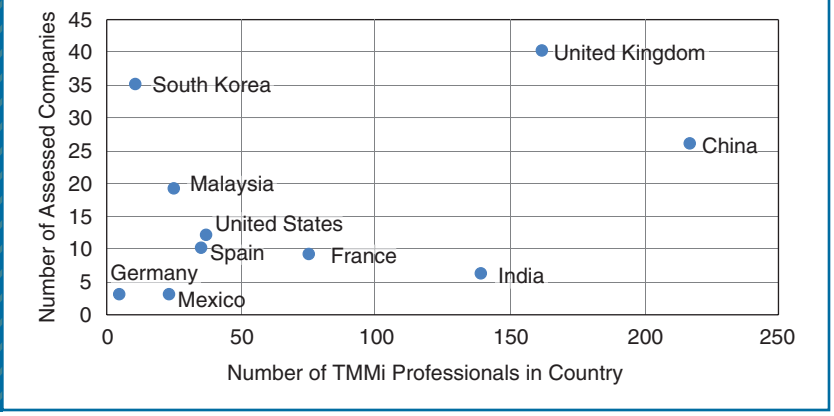


FIGURE 4. The country locations of TMMi professionals versus the number of assessed organizations.

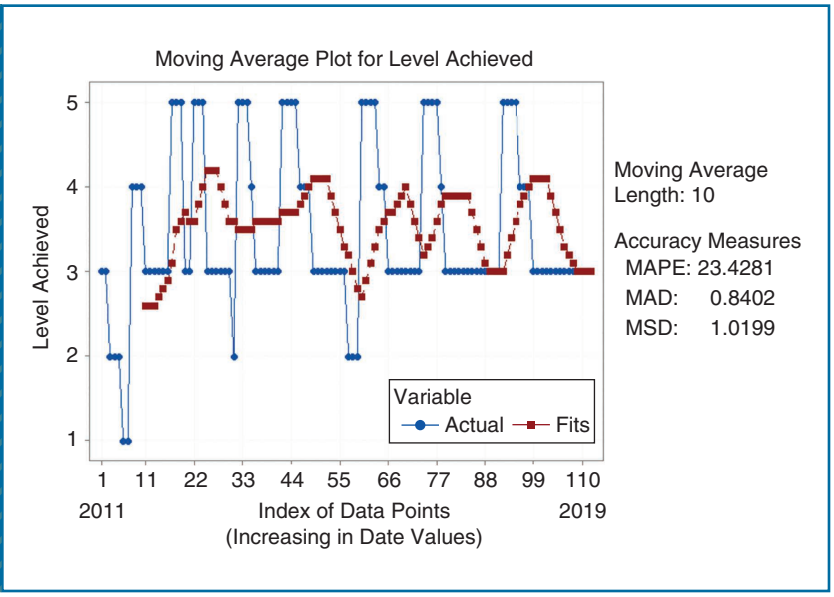


FIGURE 5. The mean (average) and median of TMMi levels achieved by certifications annually (N = 111). MAPE: mean absolute percent error; MSD: mean squared deviation; MAD: mean absolute deviation.

Score of Each PA

We also had detailed data for each PA. Let us recall from Figure 1 that, across the five levels of TMMi, there are, in total, 16 PAs. We wondered if there are any particular PAs in which companies typically were more "challenged" to pass. We should mention that, to rank the score of each PA for a given applicant (company or team), the TMMi TAMAR document[10] provides a five-point rubric (scale):

1. fully achieved (or fully implemented)
2. largely achieved
3. partially achieved
4. not achieved
5. not reviewed.

This is similar to CMMI assessments.

Figure 6 shows a stack chart for each of the 16 PAs and five scales of the scores for the 114 companies that applied for formal assessments. As we can see, most PAs were scored as "fully achieved" for most of the applicants. Some PAs were "largely achieved." A small ratio was partially or not achieved. For many of the applicants, PAs in levels four and five were "not reviewed" since, when submitting an application, a company specifies the level at which it is intending to have the assessments done. Most of the applications targeted levels two or three, and, thus, PAs in levels four and five did not have to be assessed.

A maturity level may be rated as "achieved" by the organization if all PAs in scope have been rated as either "largely achieved" or "fully achieved." Furthermore, a higher maturity level cannot be achieved without the lower

stages also being met. Thus, it is critical that a given applicant company ensures that it has the evidence/capability to "largely" or "fully" achieve (satisfy) all PAs of the level for which it intends to get certified.

Back to our question of whether there were any particular PAs in which companies typically were more challenged, we found that PA 2.5 (test environment) and PA 3.4 (nonfunctional testing) have relatively more "not achieved" scores compared to other PAs. (This is visible in Figure 6.) It seems that, for many companies in the pool of assessments, satisfying (meting) these two PAs is often challenging.

For example, it could be that some companies had not invested enough in conducting nonfunctional testing and setting up systematic test environments. The former (nonfunctional testing) is often a specific expertise that needs more resource investments by a team/company and more difficult to become mature at than, compared to other PAs. Thus, an actionable insight for teams considering applying for TMMi is to improve those aspects in their teams, before making their formal applications for TMMi.

This brief status report aimed at providing insights into the trends of worldwide TMMi assessments and certifications as well as a general picture of how companies are ranked in TMMi maturity levels. Of course, there are several other important issues related to TMMi that are worth investigating, e.g., the impacts of a TMMi certification on the quality of the software developed by an organization.

Data from a recent survey performed by the TMMi Foundation in 2020 (results are not published yet) could shed some light on this important question. Three TMMi level-three companies, which participated in the survey, self-reported defect detection percentage (DDP) improvements of 10%, 20%, and 22% (17% on average). Let us remember that DDP is

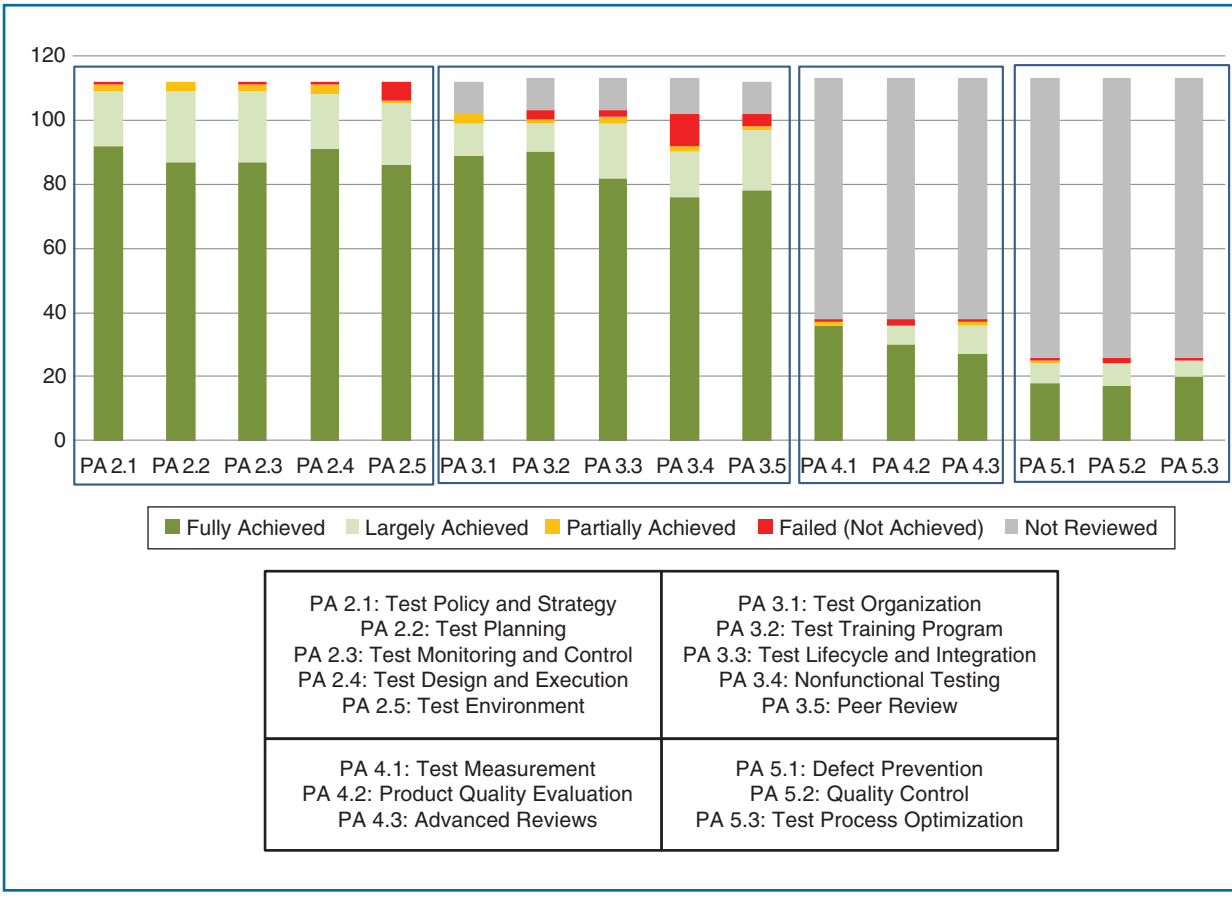


FIGURE 6. The score of each PA for the 114 companies that have undergone formal assessments.

ABOUT THE AUTHORS

VAHID GAROUSI is the CEO of Bahar Software Engineering Consulting, Belfast, U.K., and an associate professor at Queen's University Belfast, Belfast, U.K. His research interests include software testing and improving software engineering practices. Garousi received his Ph.D. in software engineering from Carleton University. Further information about him can be found at https://www.vgarousi.com. Contact him at vgarousi@gmail.com.

ERIK VAN VEENENDAAL is the CEO of the TMMi Foundation, Chester, CH1 2DS, England, and a former senior lecturer at the Eindhoven University of Technology. His research interests include software testing and test improvement. van Veenendaal received his master's degree in business economics from Tilburg University. Further information about him can be found at http://www.erikvanveenendaal.nl/. Contact him at erik@erikvanveenendaal.nl.

the number of defects found by a test phase divided by the number found by that test phase and any other means afterward. Thus, this offers some evidence of the impacts/benefits of TMMi certifications for software teams.

Despite many software quality initiatives in the last several decades, the software industry is still struggling to deliver perfect (defect-free) software. It has become apparent that, to achieve better product quality, a higher level of test maturity is required. As more awareness is raised regarding the cost of poor testing, e.g., by Britton et al.,[12] the industry is investing more resources on software testing.

Companies are finding ways to undertake testing by more effective and efficient approaches, often by conducting TPIs. TMMi is one of the established means to do so. When using TMMi, various benefits have been reported on both product quality (test effectiveness) and test efficiency,[13] reduction in test-execution times, and increased DDP.

Since there are similarities between TMMi and CMMI, it is also important to compare them. Starting in the late 1980s, CMMI has become a popular model for SPI, with a large uptake worldwide, mostly in the government and defense sectors. The fact that CMMI was initially required by the U.S. Department of Defense helped enormously with achieving its popularity. For the case of TMMi, such a driving force is slowly happening. For example, the Malaysian government issued a policy in 2018 that a company can be assigned as an independent verification and validation provider for public-sector information and communications technology sector projects only if it is at TMMi level three or above (https://bit.ly/MalaysiaTMMi).

Our analysis in this article showed that, since starting TMMi assessments in 2011, the number of annual assessments has been between 15 and 30 companies each year. In 2019, there was already a growth in the uptake of TMMi (as depicted in Figure 2).

Another interesting recent development (since 2017) for TMMi is probably the establishment of so-called TMMi local chapters. A TMMi local chapter ensures that TMMi professional training and assessment services are available locally. Already, the success from this approach is reflected in the recent growth numbers.

Furthermore, the TMMi Foundation and ISTQB, the world-leading organization for test certifications, entered into an alliance in mid-2019 to further promote the software testing profession together. The alliance aims to bring together the people (ISTQB) and process (TMMi) aspects of testing. Based on these developments, we forecast more TMMi "uptake" in the coming years as more companies see the benefits of TMMi.

References

1. V. Garousi, M. Felderer, and T. Hacaloğlu, "What we know about software test maturity and test process improvement," *IEEE Softw.*, vol. 35, no. 1, pp. 84–92, 2018. doi: 10.1109/MS.2017.4541043.
2. D. Gelperin and B. Hetzel, "The growth of software testing," *Commun. ACM*, vol. 31, no. 6, pp. 687–695, 1988. doi: 10.1145/62959.62965.
3. B. Ilene, H. Ariya, G. Robert, and C. R. Carlson, "A model to assess testing process maturity," *Crosstalk, J. Defense Softw. Eng.*, vol. 11, pp. 26–30, Nov. 1998.
4. "TMMI specification (reference model), release 1.0." TMMI Foundation. http://www.tmmi.org/pdf/TMMi.Framework.pdf (accessed Oct. 2015).
5. "Test Maturity Model integration (TMMi) reference model (specification), release 1.2." TMMI Foundation. https://www.tmmi.org/tmmi-documents/ (accessed May 2020).
6. E. v. Veenendaal, C. Shang, and Y. Xu, "Achieving TMMi Level 3–A Chinese case study," *Quality Matters Mag.*, no. 8, pp. 18–21, 2019. [Online]. Available: https://www.tmmi.org/tm6/wp

-content/uploads/2019/04/Quality_Matters-Issue-8-TMMi-case-study.pdf
7. R. Kerli and M. Raimundas, "Empirical analysis of the test maturity model integration (TMMi)," in *Information and Software Technologies*, vol. 403, T. Skersys, R. Butleris, and R. Butkiene, Eds. Berlin: Springer-Verlag, 2013, ch. 32, pp. 376–391.
8. A. Snehal and G. Kerstin, "Evaluation of test process improvement approaches: An industrial case study," Master of Science Thesis in Software Engineering, Univ. of Gothenburg, Sweden, 2013.
9. A. v. Ewijk, B. Linker, M. v. Oosterwijk, and B. Visser, *TPI Next: Business Driven Test Process Improvement*. Den Bosch, The Netherlands: UTN Publishing, 2009.
10. "TMMi Assessment Method Application Requirements (TAMAR), Version 2.0." TMMI Foundation. http://www.tmmi.org/sites/all/themes/softtech/pdf/TMMi.TAMAR.pdf (accessed Oct. 2015).
11. M. Say, "The case for TMMi." The Guardian. https://www.theguardian.com/government-computing-network/2011/sep/01/tmmi-hoit-testing (accessed Mar. 8, 2021).
12. T. Britton, L. Jeng, G. Carver, P. Cheak, and T. Katzenellenbogen, "Reversible debugging software," Judge Business School, Univ. of Cambridge, Cambridge, MA, Tech. Rep., 2013.
13. E. v. Veenendaal, "TMMi: The world standard for test process improvement," *Quality Matters Mag.*, no. 7, pp. 26–29, 2018. [Online]. Available: https://www.tmmi.org/tm6/wp-content/uploads/2018/12/QualityMatters-TMMi-the-World-Standard-for-Test-Process-Improvement.pdf

Digital Object Identifier 10.1109/MS.2022.3142496