Automated Testing and TDD

# About Automated Testing

There are many types of automated tests, but in this document, we will be focusing strictly on automated developer tests. Automated tests, as the name suggests, are a method of automatically testing and verifying the correctness of a system. These tests will return a result, which will be a success or failure based on one or many conditions.

# Benefits of Automated Testing

# Confidence

What does fear do to a software developer?

* Makes you tentative
* Makes you want to hole up and communicate less
* Makes you shy away from feedback
* Makes you unhappy

Automated testing allows you to program courageously. Move forward confidently knowing that every step of the way your software still works. We have a constant balance between fear and confidence. Anytime we become fearful, we write more tests. Once we feel we are on solid ground again. If we approach boredom, we write fewer tests. We never want to reach a point where we are writing tests for the sake of writing tests. Each test should be valuable, and be carefully evaluated from a value, cost, and risk perspective.

We can now confidently refactor legacy code knowing that if all of the tests pass that our software behaves as expected.

The below excerpts from an article written by Mike Bland, a former Googler, speaks toward this with the Google Web Server team as an example.

# The Google Web Server Story

“Despite the risks and the costs, it's important to realize that the benefits of unit testing go beyond merely minimizing the chances of releasing catastrophic bugs.

When I joined Google in 2005, it was already very successful and many "long timers" believed it was because we were doing everything right. As a result, at that time and for some years afterwards, there was a lot of resistance to change. However, as the user base and potential for catastrophe exploded, and as success and the growth that came with it caught up to Google, it became clear that more “rock stars” producing “rock star” code was going to produce nothing but a bunch of noise and confusion in the long-term. An influx of new Google developers eventually helped accelerate the cultural shift towards unit testing adoption, both because these new developers were open to the idea, and because testing eventually proved effective in helping these new folks get up to speed and avoid making mistakes.

As a concrete example, let's take what is possibly the most popular page on the Internet: Google's home page. The Google Web Server (GWS) team's unit testing story became well-known throughout the company. The GWS Team had gotten into a position in the mid 2000's where it was difficult to make changes to the web server, a C++ application serving Google's home page and many other Google web pages. Despite this difficulty, integrating new features was integral to the success of Google as a business. The barrier that was stopping people from making changes as rapidly as possible was the same that slows change on most mature codebases: a quite reasonable fear that changes will introduce bugs.

[Fear is the mind-killer.](http://en.wikipedia.org/wiki/Bene_Gesserit#Litany_against_fear) It stops new team members from changing things because they don't understand the system, and it stops experienced people changing things because they understand it all too well.

The Google Web Server Team took a hard line: No code was accepted without an accompanying unit test.

Determined to overcome this fear, the GWS Team introduced a testing culture. They took a hard line: No code was accepted, no code review was approved without an accompanying unit test. This often frustrated contributors from other teams trying to launch their features, but the GWS Team stuck to its guns.

Over time, unit test coverage and development momentum went up, while defect, production rollback, and emergency release counts went down. New team members found themselves becoming productive far more quickly because the tests allowed them to gain a deeper perspective on a system one unit at a time, and to begin contributing changes with the confidence that the existing tests would likely detect any unexpected side-effects. Any tests they caused to fail in the course of their early efforts accelerated their grasp of the system. Experienced members of the team, who had grown cautious of making changes and accepting changes from contributors, were able to make and accept changes quickly for the same reason and no longer had to rely primarily upon large and expensive system or manual tests with feedback cycles on the order of hours or days. Adding more new developers actually allowed the team to move faster and do more, avoiding the scenario described by [Brooks's Law](http://en.wikipedia.org/wiki/Brooks%27_law) in which "adding manpower to a late software project makes it later".

Furthermore, the mitigation of fear led to the expansion of their joy in programming, as they could see tangible progress being made towards exciting new milestones without being held back by chronic outbreaks of high-priority bugs. The impact on productivity of high morale, based on the ability to remain in a state of creative flow, cannot be overstated. While I was at Google, the GWS Team exhibited the ideal testing culture, integrating an enormous number of complex changes from outside contributors while making their own constant improvements.

Thanks to the GWS example inspiring the efforts of the Testing Grouplet (a team of developers volunteering to promote unit testing adoption, described in a later section of this article), many teams at Google were able to transition to a unit testing culture and benefit from reduced fear and increased productivity. It did take time to overcome inertia, indifference, the friction of outdated tools, and resistance, since at first unit testing felt like a cost and some people worried that the time spent writing that second representation of behavior could be spent writing new code (that would get them promoted). Eventually, as people experienced what it meant to cast aside the fear of change, they came to see this side-effect as easily outweighing those lines of code, in terms of its impact on their happiness, on their team's happiness, and on the bottom-line of productive output.”

<https://www.martinfowler.com/articles/testing-culture.html>

# Defect Reduction

By having a comprehensive set of tests that validate the behavior of a system, we can drastically reduce the number of defects.

“Defects destroy the trust required for effective software development. The customers need to be able to trust the software, The managers need to be able to trust the reports of progress. The programmers need to be able to trust each other. Defects destroy that trust. Without trust, people spend much of their time defending themselves against responsibility that someone else may have made a mistake”

<https://www.computer.org/csdl/magazine/so/2007/03/s3024/13rRUygT7kK>

<https://www.testingexcellence.com/pros-cons-test-driven-development/>

<https://medium.com/crowdbotics/tdd-roi-is-test-driven-development-worth-the-money-d535c8d5a5f>

<https://devblogs.microsoft.com/premier-developer/pragmatic-tdd/>

<https://martinfowler.com/articles/is-tdd-dead/>

In 2016, Google published the findings of its Aristotle Project, an investigation into what makes an effective team. The project defined what a team was, what effectiveness was, and then sought out to collect data and measure effectiveness.

The researchers identified the following dynamics, listed in order of importance:

1. Psychological Safety
   1. Team members feel safe to take risks and be vulnerable in front of each other
2. Dependability
   1. On dependable teams, [members reliably complete quality work on time](http://amj.aom.org/content/53/3/535.short) (vs the opposite - [shirking responsibilities](http://www.jstor.org/stable/pdf/258490.pdf?acceptTC=true))Structure and Clarity
3. Structure & Clarity
   1. [An individual’s understanding of job expectations, the process for fulfilling these expectations](http://www.jstor.org/stable/1556372), and the consequences of one’s performance are important for team effectiveness. Goals can be set at the individual or group level, and must be specific, challenging, and attainable. [Google often uses Objectives and Key Results (OKRs)](https://rework.withgoogle.com/guides/set-goals-with-okrs/steps/introduction/) to help set and communicate short and long term goals.
4. Meaning
   1. Finding a sense of purpose in either the work itself or the output is important for team effectiveness. The meaning of work is personal and can vary: financial security, supporting family, helping the team succeed, or self-expression for each individual, for example.
5. Impact
   1. The results of one’s work, [the subjective judgement that your work is making a difference, is important for teams](http://www.ncbi.nlm.nih.gov/pubmed/18211139). Seeing that one’s work is contributing to the organization’s goals can help reveal impact.

**[Google2016]**

Automated testing doesn’t hit on all of these points, but it does hit on the top two in a large way. By writing valuable tests that validate the behavior of our software, we have a net of psychological safety. By having a team that exercises this discipline, we can all depend on one another regarding the structure, correctness, and quality of our software.

# Improved Work Practices

Having fewer defects really empowers a team to take new and interesting approaches. We can now:

* If the defect density can be reduced enough, then quality assurance can shift from reactive work to proactive work
* If the defect density can be reduced enough, we can have shippable software with new functionality every day, leading to new business relationships with customers
* When defects are reduced to a low number, they can also be evaluated thoroughly. Why did the defect occur? Was it a misinterpreted requirement? Was there a complex implementation and an edge case was mixed? Understanding why a defect exists can be used to drive continuous improvement so that the team does not introduce similar defects in the future. When there are too many defects this is not possible, as the team is struggling just to resolve all of the defects and survive.

# Self Testing Code

<https://www.martinfowler.com/bliki/SelfTestingCode.html>

# Quick Release Cycles

In a fast moving industry where services are deployed with new functionality multiple times a day, quality checks have to exist to ensure that faulty components are not being deployed. These quality checks cannot appear in the form of mandatory gate keepers (Certification and Delivery), or scheduled code reviews, since those will slow the process down too much. Tests fill this void.

# Clean, Decoupled Code

Writing tests for code that is tightly coupled, is very difficult, and can be entirely impossible if you strive for testing only one unit at a time. TDD naturally supports software design principles such as the Open / Closed Principle, Dependency Inversion Principle, and Single Responsibility Principle. You can even build a dynamic sweet of unit tests to run against all classes that implement a given interface and evaluate if it violates the Liskov Substitution Principle.

# Executable Documentation

Well-written unit tests can provide two types of documentation: the test names act as a sort of specification of the code's behavior; and the tests themselves act as code samples for each behavior case. Even better than typical Application Programming Interface (API) documentation, well-maintained unit tests are by definition an up-to-date representation of actual behavior. The author of a unit test effectively communicates to other developers how a piece of code should be used, and what to expect from it. These "other developers" may be brand new to the team, or may not yet be hired (or even born). Such documentation helps developers understand unfamiliar code, even entire systems, without interrupting anyone else to the degree that they might without unit tests.

Poorly-written unit tests lack this quality, usually because less thought is given to test code than "production" code. The solution: Set the same quality bar for test code as production code. If you don't, your tests will become hard to maintain and slow down the team.”

<https://www.martinfowler.com/articles/testing-culture.html>

# Accelerated Understanding

“Think of it like this: Every time a test fails, that is an opportunity to deepen your understanding of the system. If you're new to a team, breaking many tests as you begin to make changes to the system can help you become productive far more quickly, as each of these events align your awareness of the system more closely with reality. If you've been on the team for a long time, existing tests will answer many questions that new contributors may have, saving your time and focus. They will also remind you of all the nuances of the code you might have written in the past, and haven't had to think about for some time, should you have to dive back into it. In other words, you benefit your future self when adding a well-crafted suite of tests to your code, minimizing the time needed to context-switch back into that prior state of mind.

Think of the opposite, as was the case in the pre-unit testing days of GWS: When you're on a project that doesn't have ample unit testing coverage, you're afraid to do anything since you don't know what you might break.”

<https://www.martinfowler.com/articles/testing-culture.html>

# Faster Bug Hunting

“Imagine a bug is found in integration or system testing, or after a new release is pushed to a datacenter, or perhaps by a user some time after that. The developers responsible for the buggy code have already moved on to other tasks, and are likely under deadline pressure to deliver. If the bug is severe enough, at least one of those developers will have to stop to address it, slowing the progress of the new development work underway.

If the buggy code is well-covered by a suite of automated tests, especially small unit tests, this interruption may not take much time on the part of the developer assigned to fix the bug. The existing tests serve as documentation of the intent of the affected code. The developer adds a new test to reproduce the bug, verifying that the defect is well-understood before attempting to fix it. This new test verifies the fix for the bug, and the existing tests provide a high degree of confidence that the fix is free of unintentional side-effects. The new test becomes a permanent part of the test suite to guard against regression, the fix is released, and development on the new release continues. The interruption is finished.

Contrast that against the situation where the buggy code isn't well-covered by unit tests. The developer must take time to understand the affected code and far more care to pinpoint the error and ensure its fix is free of side-effects. Verification of the fix may not come for days or even longer, depending on the nature of whatever pre-release testing happens to be in place, if any. The interruption is prolonged, and drains more development and testing time from the new release.

Or, even worse: The team may decide to leave the bug in-place from fear of breaking something else. That certainly doesn't inspire user trust, much less developer confidence and productivity.”

<https://www.martinfowler.com/articles/testing-culture.html>

# About TDD

### What is test driven development?

Test Driven Development is a technique that involves writing tests prior to writing the associated implementation. There are various rules that are followed to various levels, such as:

* Write new code only if an automated test has failed
* Eliminate duplication

### Red, Green, Refactor

The primary mantra of TDD is Red, Green, Refactor. This involves:

* Writing a failing test that will not compile. (Red)
* Writing enough code to make your test compile, but not pass.
* Writing enough code to make your test pass. (Green)
* Refactoring your code to be cleaner and eliminate duplication (Refactor)

# **Benefits of TDD**

# TDD Helps Decompose Problems

“Imagine programming as turning a bucket of water from a well. When a bucket is small, a free spinning crank is fine. When a bucket is big and full of water, you’re going to get tired before the bucket is all the way up. You need a ratchet mechanism to enable you to rest between bouts of cranking. The heavier the bucket, the closer the teeth on the ratchet need to be. The tests in test driven development are the teeth on the ratchet.”

TDD is a tool to manage difficult tasks, one small step at a time.

“Unit testing is not in the same class as integration testing, or system testing, or any kind of adversarial "black-box" testing that tries to exercise a system based solely on its interface contract. These types of tests can be automated in the same style as unit tests, perhaps even using the same tools and frameworks, and that's a good thing. However, unit tests codify the intent of a specific low-level unit of code. They are focused, and they are fast. When an automated test breaks during development, the responsible code change is rapidly identified and addressed.

This [rapid feedback cycle](http://en.wikipedia.org/wiki/OODA_loop) generates a sense of [flow](http://en.wikipedia.org/wiki/Flow_(psychology)) during development, which is the ideal state of focus and motivation needed to solve complex problems. Contrast that with the opposite phenomenon, using the familiar operating systems metaphor of [context switching](http://en.wikipedia.org/wiki/Context_switch). Context switching requires that the present state of operations be saved somehow, and that a new state of operations be swapped in before initiating the new activity; then there's the time and effort involved in switching back. Plus, there's the issue of how much state must be managed per operation. Without unit tests, we have to use more of our brains to remember weird corner cases and strange side-effects, giving us less time and energy to do the thing we're better at than the computer: advancing solutions to new problems rather than juggling the weight of all the problems that have already been solved.

In other words, you can be more productive since you can iterate on code much quicker: You don't need to start up some heavyweight server if you can just run a unit test instead. So if it takes a few tries to get some code right, those few tries might take minutes (or longer) if you have to start up a server again and again, compared to seconds if you just need to rerun the unit tests each time.”

<https://www.martinfowler.com/articles/testing-culture.html>

# Improved Code Qualilty

Since you write tests firsts with TDD, you discover coupling between classes quickly. This becomes less of an issue as developers become more experienced and learn to design effective code without the guardrails of TDD, but when starting out, tightly coupled classes is a major concern.

“Far from being an exercise in academic purity, code quality matters. Bad code provides bugs with plenty of shadows in which to hide; good code increases the chances that they will be found and squashed sooner rather than later. When the author of a piece of code writes a test for that code, the author effectively becomes the first user. Just as [eating your own dogfood](http://en.wikipedia.org/wiki/Eating_your_own_dog_food) is good software development practice at the overall product level, having to write code that uses your own code can lead to improved designs that are more readable, maintainable, and debuggable.

Think of what problems you're trying to solve with the code you're writing; then think of the code you'd like to write, as a client, to make use of the solution. That ideal client code can be expressed as unit test cases that use the interface of the code you're developing.

When code-level design is approached this way, all of the smaller pieces that make up the larger system become not just more reliable, but easier to understand. This makes everyone more productive, as the mental effort required to comprehend what a specific piece of code does is minimized.”

<https://www.martinfowler.com/articles/testing-culture.html>

# TDD Evaluates Assumptions Early

TDD forces you to evaluate your assumptions about how the system should work immediately. You state your test cases up front, and then design a solution that will satisfy that criteria.

When writing tests gets challenging, it is often a sign that our software is becoming rigid. That being said this is not always the case, and it should be carefully evaluated. Changing software for the sole purpose of allowing a component to be tested (when the change otherwise is not a valuable change) is often considered a code smell.

# TDD is Satisfying

TDD is a stress reliever. It feels good to make predictable, measured progress. It feels good to know that each step of the way all of the previous functionality is still working as expected. It is extremely satisfying to write a new test, and after implementing the required functionality have it flash green. It turns software development into a type a minigame that I’m playing. All of these things make a life as a software developer more enjoyable.

# **Automated Testing Trade Offs**

# Developers Must Be Trained in Automated Testing

“There will be a learning curve. Like any skill that relies on craft rather than rote processes, a programmer learning to write unit tests will have to go through phases of learning and development, of trial and error, reflection, experimentation, and integration. This takes time, energy, and funding away from other activities. It will cause an initial slow-down in development as people grow accustomed to the practice.

That said, this is a one-time cost. The cost of bringing someone up to speed on unit testing is relatively low if good unit testing practices are already in place on a team, and unit testing skills are portable from one project to the next. Hence, the learning curve is steepest for teams that don't have any unit testing practices at all.

Unit testing, like any other tool, language, or process, can be applied poorly—especially when one first begins, and even more so if one has no good examples to follow, nor mentors for guidance. Unit tests which are brittle, large, slow, perpetually broken (and subsequently ignored), or flaky set bad examples which can get replicated through an entire test suite like a virus. Poorly-written tests can actually be worse than no tests at all, leaving the impression that testing is a waste of time. Builds remain broken and ignored, flooding the testing signal with the noise of constant failures. Developers uninterested in working with the testing environment become willing to live with the fear of making slow and painful changes. The end result is a drag on productivity, an increased risk of defects, and a team convinced that testing is for other people.”

“To remedy this lack of knowledge and experience, motivated developers can band together to improve one another's unit testing skills and increase the amount of test coverage of the code base over time. In this section, I will describe how the Google Web Server Team built up its test coverage and achieved a high degree of overall productivity; in later sections I will explain how Google as a whole was able to adopt a unit testing culture, and how lessons from that experience may apply to individual teams. However, self-training will take time and energy, and the big-picture payoff may not be immediately apparent, so it requires patience, honest effort, and commitment to see all the way through. Over time, though, as the code base grows and more developers join the team, the value becomes increasingly clear. A two-person team might manage without unit testing, but a twenty-person team will have a harder time, as feature and communication complexity is compounded.

If developers are not motivated to research available materials and improve their skills, or just don't have any idea how to begin, this may imply the need to invest in internal training programs or to contract outside help to provide training. This can lead to a bit of price-shock if resources are tight, deadlines are looming, and the future benefits do not seem clear. The time required to learn the necessary skills should be no greater than that required to train developers in any other skill or technology; but if developers resist, the process can become more drawn-out, painful, and expensive.”

<https://www.martinfowler.com/articles/testing-culture.html>

# Tests Must be Maintained

Tests should be treated as first class citizens. They are very important and must be maintained properly. If you developers cannot trust the tests in terms of reliability or if the tests take too long to run, the developers will no longer use the tests. If your team has doubts about what the tests say about the reliability of the code, you lose many of the advantages that automated testing provides.

“Sometimes, tests themselves can become a maintenance burden; it may seem like they paint a project into a corner, restricting progress rather than maximizing it. This is a particular danger to new teams that lack experience with unit testing and don't understand its value. Mock objects are prone to misuse by inexperienced practitioners, leading to brittle tests of dubious value. With experience, this scenario becomes less likely. You eventually learn to step back, reevaluate the goal of the code and the test, and rewrite one, the other, or both. In the meanwhile, it may become necessary at times to replace an overly-restrictive test rather than to spend the effort salvaging it.

Speaking of new projects or teams or companies or domains, as ideal as it may be to follow [Agile](https://www.martinfowler.com/agile.html) practices to the letter and practice pure [Test Driven Development](https://www.martinfowler.com/bliki/TestDrivenDevelopment.html) (TDD) at all times, sometimes a developer or a team needs to explore, to play, before getting serious about defining expectations and behavior. (Some argue that always following all Agile practices to the letter is a demonstration that you don't understand Agile.) While it’s always nice to get testing experience as early as possible on a project, sometimes you just need to write throw-away, prototype code; in that case thorough unit testing is probably overkill. This may be especially true of startups trying to launch a product as fast as possible.

On the other hand, be aware of the saying: "There is nothing more permanent than throw-away code." The trade-off is that the more features are implemented without accompanying tests, the more [Technical Debt](https://www.martinfowler.com/bliki/TechnicalDebt.html) a team builds up that must be repaid later. Unit testing can be difficult if you don't design for testability from the start—using [dependency injection](https://www.martinfowler.com/articles/injection.html), writing well-defined classes that focus on one thing, and so forth. It is up to the team to gauge the acceptable limits of such debt, and at which point it must be paid to avoid an even more expensive rewrite once maintenance and new feature development grow too cumbersome.”

<https://www.martinfowler.com/articles/testing-culture.html>

# Tests Can Result in a False Sense of Security

“There's no guarantee that unit tests themselves will be bug-free. Consider this example (in C++-like pseudocode based on the [Google Test](https://code.google.com/p/googletest/) framework):

TEST\_F(FooTest, IfAPresentFilterB) {

setup input and add "A:" , "B:"

run call

EXPECT\_TRUE(PresentInOutput("A:"))

EXPECT\_FALSE(PresentInOutput("B"))

}

The second expectation in this test should check for "B:", with a colon, not just "B". If the code under test accidentally filters for "B" without a colon, the test will pass when it should fail.

It's arguable that the test makes things worse in this case, providing a false sense of security. However, the bug could exist even if the test hadn't been written; given the existence of the buggy test, fixing the code and the test is tantamount to providing a regression test for the bug. Fixing the test and learning from the mistake provides value; blaming the test and deleting it is a step backwards. As one possible measure to avoid buggy tests in the future, the team responsible for such a bug could endeavor to take a closer look at the test code submitted as part of future code reviews, to provide it with the same priority and care as "production" code.

In practice, buggy unit tests tend to be the exception. If practicing pure Test-Driven Development, a failing test should be written before the code that makes it pass; this could help to prevent such bugs. If not practicing pure TDD, temporarily adding an error into the code under test to make sure the test will fail can also help. In either case, writing multiple test cases that check that the code doesn't do what it shouldn't do (instead of just checking the happy path where all inputs are valid) may reveal bugs in other test cases. Still, the possibility remains that unit tests themselves may contain bugs, especially if care isn't taken to ensure that they fail when they're supposed to.”

# Test Induced Design Damage

Sometimes software designs are not easily testable. Abstractions may be introduced for the sake of testing, and only for the sake of testing. These abstractions are test induced design damage. Generally, modifying code for the sole purpose of testing is considered a code smell. That being said, there are times where having the test can be very valuable. You must evaluate how valuable the test you are writing is and contrast that value with the potential damage being done to your code base. In many cases, there will be no damage at all. However, in cases that require mocking, indirection and abstraction for the sake of testing, carefully consider the damage that is being done to the cohesion and readability of the code for their sake.

Arguments are often made that even if an abstraction is only created for the use of tests, it does keep options open in the future for alternate implementations. This is often a defense given for creating an abstraction over the data access components. It is also correctly pointed out that these implementations rarely change. You are unlikely to swap from SQL Server to Oracle in a project. While “swapability” can be valuable, it is important to be honest with ourselves about the potential of a swap to actually happen.

# **TDD Trade Offs**

TDD has become dogmatic and moralistic

### Encourages Testing Extremes

TDD has been criticized for encouraging testing extremes that are not necessarily the correct approach in all cases. If I am creating software that will be utilized by a pacemaker, and will responsible for keeping people alive, my unit testing coverage and reliability of my software will be held to a much higher standard than that of a CRUD application. Achieving higher levels of reliability requires an exponential increase in costs. 99% reliability is several times more expensive than 95% reliability, and 99.99999% reliability is several orders of magnitude more expensive again.

TDD can take a hard line that 100% test coverage is a goal. It is often stated with TDD that you should not write any production code without it being mandated by a failing test. These are extremes, and are not necessarily valuable. In fact, these extremes can be very damaging. We should make cost/benefit decisions, and never resort to a default answer based on a dogmatic rule.

<https://dhh.dk/2014/tdd-is-dead-long-live-testing.html>

<https://martinfowler.com/articles/is-tdd-dead/>

The above link contains about 2 ½ hours of discussion between David Heinemeier Hansson, Martin Fowler, and Kent Beck on TDD. I highly recommend giving it a watch. It challenges several assumptions about TDD and has some great value based discussions about TDD and automated testing.

### Refactoring Can Be Skipped

TDD is driven by the Red, Green, Refactor loop, but the dopamine spike driven by the green bar arrives during the Green phase. At times, the refactor step is skipped resulting in suboptimal code.

### TDD is not the way everyone’s brain works

For some people, it is very natural to write a test first and then the create an implementation that satisfies the given condition. For others, they think about this in the inverse. It is difficult to think about what the test should be without

### TDD can result in

As humans it is tempting to try and quantify things. Quantifying the quality of a design is really difficult. We may tend to focus on things that are easily measurable but less important than things that are more important but not as easily measured.

Make certain to constantly evaluate trade offs

### TDD is not how everyone’s brain works

The ability to write tests for anything is a valuable skill to have. It is the four wheel drive of programming. If I get into a really sticky spot, I can always ratchet up the tests.

At the end of the day, the principles and practices and decisions are up to the team

# **Types of Automated Tests**

## Unit Tests

* A small, fast test designed to cover a single unit of work, and make a single assertion about how that unit of work behaves. These tests should make up the bulk of your testing suite. These tests should run in memory, and not require any dependencies to function.

## Integration Tests

* A test verifies if “independently developed units of software work correctly when they are connected to each other”.[1](#IntegrationTest)It is a slower test that covers a larger surface area. The size of this surface area can vary greatly. Since these tests are generally more involved and more difficult to create, there will generally be less Integration tests than Unit tests.
* <https://martinfowler.com/bliki/IntegrationTest.html>

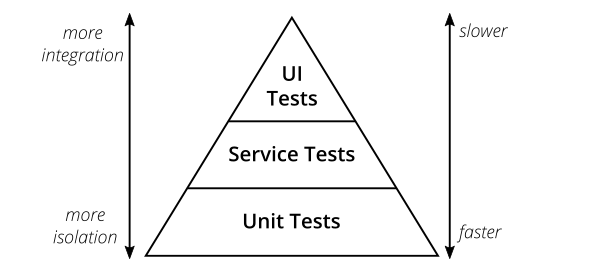
## End to End Tests

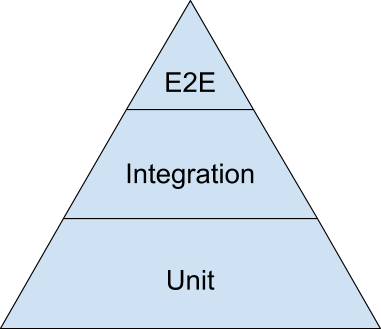
* A test that covers large sections of the system, form the start point to the end point. You ensure that your entire application and its dependent systems integrate correctly. They closely mimic how a real user would use the application.

## UI Tests

* UI testing involves verifying that UI elements perform correct behaviors. Since UIs change frequently, these tests can be expensive to maintain. A test suite will usually have less UI tests than integration tests.

You will often see these tests displayed as a Test Pyramid, representing the difficulty of writing and maintaining the tests to how many of the tests are contained within a test suite. These pyramids can differ in appearance, but the general idea is to write less tests that take a long time to run and cover large surface areas, or tests which are likely to change frequently (Such as UI tests).





There are more types of tests, such as Contract Tests, End to End tests, Spike tests, and Exploratory tests. These tests are not within the scope of this document.

It is important to keep to this type of testing structure. While end the end tests can be valuable, they are often flaky. Anything going wrong along a very large path can cause the test to fail. Many times the failure will not necessarily be a sign that the code is not working correctly. When an end to end test does fail, it can be difficult to ascertain the reason why. You must search the entire area of code covered by the test to find the failure.

<https://testing.googleblog.com/2015/04/just-say-no-to-more-end-to-end-tests.html>

This is in stark contrast to unit tests, which only cover a very specific and small section of code. If my unit test fails, I should immediately know the affected area, and why the failure occurred.

All of these types of tests exist on a continuum. There is not a clearly defined point where a unit test becomes an integration test and a integration test becomes an end to end test. That being said there are clear examples of each.

* If a test only interacts with a single class, runs exclusively in memory, and stubs and mocks dependencies, the test is clearly a unit test
* If a test tests the interaction between two classes and how they interact with the file system through an interface, it is an integration test.
* If a test covers an entry point in the UI and goes all the way to the database and back again, it is an end to end test.

Even when the lines get a little blurry between these test types, it is still important to roughly stick to a pyramid type structure. If you test suite becomes slow, or many of your tests are flaky and fail for suspect reasons, it may be a sign that you need to take a look at the composition of your test suite.

You may choose the break these tests up and only run certain tests at certain times depending on various factors. For example, since unit tests are very fast, you may choose to separate them from the integration tests and end to end tests. Developers can run all the unit tests as they develop and get quick results back. You may choose to only run end to end tests when a check in is made. Regardless, it is important that your test suite have some performance considerations. If you unit tests become too slow, your developers will stop using it as a tool.

<https://martinfowler.com/articles/practical-test-pyramid.html>

https://www.youtube.com/watch?v=wEhu57pih5w

# Unit Testing

Unit tests follow a generic pattern. Arrange, Act, and Assert

Given When Then

See the test fail, even if you are not using TDD

Keep Unit tests small

Keep tests readable

Unit tests are generally divided into two categories. State verification and behavior verification.

Tend to stick to the side of state based verification when possible.

# Mocks and Stubs

Mocking and stubbing are tools that are used to remove dependencies from tests.

**Mock:** An implementation of an interface or base class that has some predefined behavior

<https://martinfowler.com/articles/mocksArentStubs.html>

Test Coverage

<https://www.martinfowler.com/bliki/TestCoverage.html>

<http://www.exampler.com/testing-com/writings/coverage.pdf>

<http://www.developertesting.com/archives/month200705/20070504-000425.html>

Test Certified

<https://mike-bland.com/2011/10/18/test-certified.html>

# References

1. Beck, Kent, and Cynthia Andres. *Extreme Programming Explained: Second Edition, Embrace Change*. Addison-Wesley, 2015
2. Beck, Kent. *Test-Driven Development by Example*. Addison-Wesley, 2014.
3. “Re:Work.” *Google*, Google, rework.withgoogle.com/print/guides/5721312655835136/.

Resources

**Skillport:**

[**https://cgi.skillport.com/skillportfe/main.action#summary/COURSES/CDE$130120:\_ss\_cca:sd\_adev\_a07\_it\_enus**](https://cgi.skillport.com/skillportfe/main.action#summary/COURSES/CDE$130120:_ss_cca:sd_adev_a07_it_enus)

[**https://cgi.skillport.com/skillportfe/main.action#summary/COURSES/CDE$121695:\_ss\_cca:wd\_dmvc\_a04\_it\_enus**](https://cgi.skillport.com/skillportfe/main.action#summary/COURSES/CDE$121695:_ss_cca:wd_dmvc_a04_it_enus)

**Pluralsight:**

<https://app.pluralsight.com/library/courses/play-by-play-wilson-tdd/table-of-contents>

<https://app.pluralsight.com/library/courses/pragmatic-unit-testing/table-of-contents>

<https://app.pluralsight.com/library/courses/writing-highly-maintainable-unit-tests/table-of-contents>

**Books**

TDD By Example

The Art of Unit Testing