

## Chapter 10

# ACCEPTANCE TESTING

---

*What's done, is done.*

—Shakespeare, *Macbeth*, Act 3, scene 2

*If it isn't tested, it doesn't exist.*

—Anonymous agile master

*We recently transitioned to agile. But all our testers quit.*

—Vignette from Crispin and Gregory [2009]

## WHY WRITE ABOUT TESTING IN AN AGILE REQUIREMENTS BOOK?

As a sanity check in preparing for this chapter, I went to my bookshelf and looked at a number of texts on software requirements management, including my own [Lef-fingwell and Widrig 2003], for guidance on testing whether an application meets its requirements. Of course, I knew I wouldn't find much on testing there, if for no other reason than I knew I hadn't written much. I wasn't surprised that other requirements authors haven't written much of anything on testing either.

So, the question naturally arises: Why do we feel compelled to write about testing now, in a book on agile requirements? The question itself reflects a traditional view, that historically, software requirements were somehow independent of their implementation. They lived a separate life—you could get them (reasonably right) at some point, mostly up front; the developers could actually implement them as intended; they would be tested somewhat independently to assure the system worked as intended; and the customers and users would be happy with the result. Of course, it never really worked that way, but it sure was easier to write about it.

In thinking in lean and agile terms, however, we must take a much more systemic and holistic view. We understand that stories (requirements), implementation (code), and validation (acceptance tests, unit tests, and others) are not separate

activities but a continuous refinement of a much deeper understanding; therefore, our thinking is different:

No matter what we thought earlier in the project, *this* functionality is what the user really needs, and it's now implemented, working, and tested in accordance with the continuous discussions and agreements we have forged during development.

Just as importantly, we have instrumented the system (with automated regression tests) such that we can assure this functionality will continue to work as we make future changes and enhancements to the system. Then, and only then, can we declare that our work is complete for this increment.

That is the reason that we have taken a much more systemic view of “requirements” in this book—discussing users, agile teams, agile process, roles, product owners, and whatever else is necessary for a team to develop an application that, in the end, actually solves the user's problem.

However, when describing requirements in book form, the subject is fuzzier and more tangible at the same time. It's *fuzzier* because you can't really tell where a story ends and its acceptance test begins. Are the data elements in a user entry field included in the story, or are they implied requirements attached to the story? Are they really details left for the acceptance test? Or are they perhaps so fine-grained that they may be covered solely in the unit tests for the method that implements it?

And yet, the subject is *more tangible*, because the precise answers to these questions aren't so important. What is important is that we worked through a cycle of incremental information discovery; we collaborated, we negotiated, we refined, we compromised, and we ended up with something that actually works. In addition, we have captured the details of system behavior in a set of tests that will persist for all the time the software continues to provide value to its users. So, the requirements are implemented, complete, and tested.

In this chapter, we'll provide guidance to these questions and an overview of how we achieve quality in our agile requirements practices. In agile, we simply can't do that without a discussion of testing.

## AGILE TESTING OVERVIEW

Given that we are taking a systemic view to requirements, across the team, program, and portfolio, we must also take a broader view of agile testing in general, so we'll know the context in which a discussion of acceptance testing can make sense.

Brian Marick, an early XP proponent (and a signer of the Agile Manifesto), has provided much of the thought leadership in this area and has developed a framework that many agilists use to think about testing in an agile paradigm. His philosophy of agile testing is as follows:<sup>1</sup>

Agile testing is a style of testing, one with lessened reliance on documentation, increased acceptance of change, and the notion that a project is an ongoing conversation about quality.

He goes on to describe two main categories of testing: *business-facing* and *technology-facing* tests:<sup>2</sup>

A **business-facing** test is one you could describe to a business expert in terms that would (or should) interest her .... You use words drawn from the business domain: “If you withdraw more money than you have in your account, does the system automatically extend you a loan?”

A **technology-facing** test is one you describe with words drawn from the domain of the programmers: “Different browsers implement JavaScript differently, so we test whether our product works with the most important ones.”

He further categorizes tests, whether business-facing or technology-facing, as being used primarily to either *support programming* or to *critique the product*.

Tests that **support programming** mean that the programmers use them as an integral part ... of programming. For example, some programmers write a test to tell them what code to write next .... Running the test after the [code] change reassures them that they changed what they wanted. Running all the other tests reassures them that they didn't change behavior they intended to leave alone.

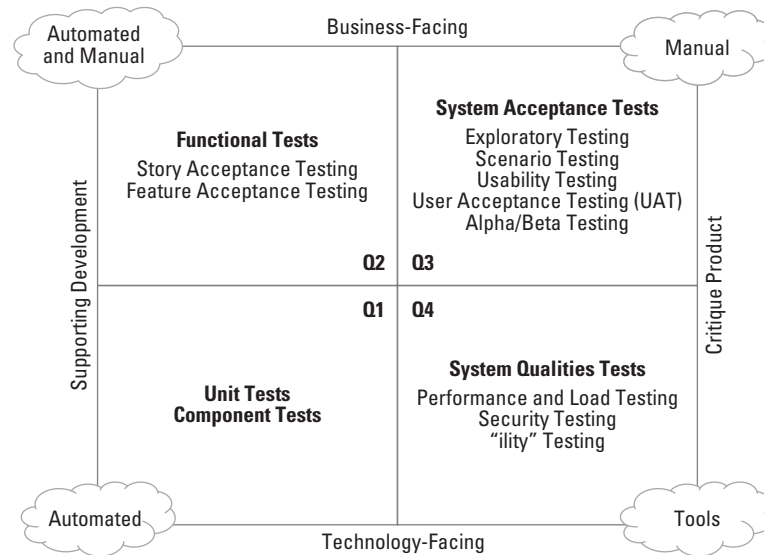
Tests that **critique the product** are not focused on the act of programming. Instead, they look at a finished product with the intent of discovering inadequacies.

In *Agile Testing*, Crispin and Gregory developed these concepts further [2009]. With a few minor adaptations for our context, we find an agile testing matrix in Figure 10–1.

---

1. [agilemanifesto.org/authors.html](http://agilemanifesto.org/authors.html)

2. [www.exampler.com/old-blog/2003/08/21/](http://www.exampler.com/old-blog/2003/08/21/)



**Figure 10–1** The agile testing matrix

In quadrant 1, we find *unit tests* and *component tests*, which are the tests written by developers to test whether the system does what they intended it to do. As indicated on the matrix, these tests are primarily *automated*, because there will be a very large number of them, and they can be implemented in the unit testing environment of choice.

In quadrant 2, we find *functional tests*. In our case, these consist primarily of the story-level acceptance tests that the teams use to validate that each new story works the way the product owner (customer, user) intended. Feature-level acceptance testing is referenced in this quadrant as well. Many of these tests can be automated, as we'll see later, but some of these tests are likely to be manual.

In quadrant 3, we find *system acceptance tests*, which are system-level tests to determine whether the aggregate behavior of the system meets its usability and functionality requirements, including the many variations (scenarios) that may be encountered in actual use. These tests are largely manual in nature, because they involve users and testers using the system in actual or simulated deployment and usage scenarios.

In quadrant 4, we find *system qualities tests*, which are used to determine whether the system meets its nonfunctional requirements. Such tests are typically supported

by a class of testing tools, such as load and performance testing tools, which are designed specifically for this purpose.

With this perspective, we have a way to think about the different types of testing we'll need to do to assure that a system performs as expected.

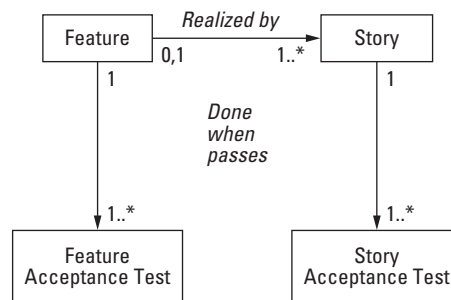
## WHAT IS ACCEPTANCE TESTING?

The language around testing is as overloaded as any other domain in software development, so the words *acceptance testing* mean different things to different people. Indeed, there are two different uses of the term in the agile testing matrix. In Figure 10–1, quadrant 2, we see *functional* and *story and feature acceptance tests*, and in quadrant 3, we see *system-level acceptance tests*. In this chapter, we'll focus on two of these quadrants, Q2 and Q1. In Chapter 17, Nonfunctional Requirements, we'll revisit the testing process to look at testing practices that support the tests indicated in quadrants 3 and 4.

In quadrant 2, we find both feature and story acceptance tests, which are used to assure that features and stories, respectively, are *done*, as we illustrate in our model in Figure 10–2.

### Story Acceptance Tests

The majority of the testing work done by agile teams is in the development, execution, and regression testing of story acceptance tests (SATs), so we will focus on those first. Story acceptance tests are functional tests intended to assure that the implementation of each new user story (or other story type) delivers the intended



**Figure 10–2** Features and stories cannot be considered done until they pass one or more acceptance tests.

behavior. If all the new stories work as intended, then each new increment of software is delivering value and provides assurances that the project is progressing in a way that will ultimately satisfy the needs of the users and business owners. Generally, the following is true.

- They are written in the language of the business domain (they are business-facing tests from quadrant 2).
- They are developed in a conversation between the developers, testers, and product owner.
- Although anyone can write tests, the product owner, as business owner/customer proxy, is the primary owner of the tests.
- They are black-box tests in that they verify only that the outputs of the system meet the conditions of satisfaction, without concern for how the result is achieved.
- They are implemented during the course of the iteration in which the story itself is implemented.

This means that new acceptance tests are developed for every new story. If a story does not pass its test, the teams get no credit for the story, and the story is carried over into the next iteration, where the code or the test, or both, are reworked until the test passes.

## CHARACTERISTICS OF GOOD STORY ACCEPTANCE TESTS

If stories are the workhorse of agile development—the key proxy artifact that carries the value stream to the customer—then story acceptance tests are the workhorse of agile testing, so teams spend much time defining, refining, and negotiating the details of these tests. That’s because, in the end, *it is the details of these tests that define the final, agreed-to behavior of the system*. Therefore, writing good story acceptance tests is a prime factor in delivering a quality system. Good story acceptance tests exhibit the characteristics described in the following sections.

### They Test Good User Stories

An attribute of a good SAT is that it is associated with a good user story. In Chapter 6, we described the INVEST model for user stories and the quality of our acceptance tests are fully dependent on the native quality of the story itself. In particular, the user story to which a test is associated has to be independent, small, and testable. If the development of an acceptance test illustrates that the story is

otherwise, then the story itself must be refactored until it meets these criteria. To get a good acceptance test, we may need to refactor the story first. For example, the following:

*As a consumer, I am always aware of my current energy costs.*

becomes this:

*As a consumer, I always see current energy pricing reflected on my portal and on-premise devices so that I know that my energy usage costs are accurate and reflect any utility pricing changes.*

### **They Are Relatively Unambiguous and Test All the Scenarios**

Since the story itself is a lightweight (even potentially throwaway) expression of intent, the acceptance test carries the detailed requirements for the story, now and into the future. As such, there can be little ambiguity about the details in the story acceptance test. In addition, the acceptance test must test all the scenarios implied by the story. Otherwise, the team won't know when the story is sufficiently complete in order to be able to be presented to the product owner for acceptance. Here's an example:

Story:

*As a consumer, I always see current energy pricing reflected on my portal and on-premise devices so that I know that my energy usage costs are accurate and reflect any utility pricing changes.*

Acceptance test:

1. Verify the current pricing is always used and the calculated numbers are displayed correctly on the portal and each on-premise device (see attachment for formats).
2. Verify the pricing and the calculated numbers are updated correctly when the price changes.
3. Verify the "current price" field itself is updated according to the scheduled time.
4. Verify the info/error messages when there is a fault in the pricing (see approved error messages attached).

### They Persist

One of the mysteries about agile, and indeed a key impediment to adoption, is a commonsense question: “If developers don’t document much and there are no software requirements specifications as such, how are we supposed to keep track of what the system actually does? After all, we are the ones responsible for assuring that it actually works, now and in the future. Isn’t that something we have to know, not just once, but in perpetuity?”

The answer is yes, indeed, we do have to know how it works, and we have to routinely regression test it to make sure it continues to work. We do that primarily by persisting and automating (wherever possible) acceptance tests and unit tests (discussed later).

User stories can be safely thrown away after implementation. That keeps them lightweight, keeps them team friendly, and fosters negotiation, but acceptance tests persist for the life of the application. We have to know that the current price field didn’t just get updated once when we tested it but that it gets updated every time the price changes, even when the application itself has been modified.

### ACCEPTANCE TEST-DRIVEN DEVELOPMENT

Beck [2003] and others have defined a set of XP practices for agility described under the umbrella label of test-driven development (TDD). In TDD, the focus is on writing the unit test before writing the code. For many, TDD is an assumed part of agile development and is straightforward in principle.

1. Write the test first. Writing the test first forces the developer to understand the required behavior of the new code.
2. Run the test, and watch it fail. Because there is as yet no code to be tested, this may seem silly initially, but this accomplishes two useful objectives: it tests the test itself and any test harnesses that hold the test in place, and it illustrates how the system will fail if the code is incorrect.
3. Write the minimum amount of code that is necessary to pass the test. If the test fails, rework the code or the test as necessary until a module is created that routinely passes the test.

In XP, this practice was primarily designed to operate in the context of unit tests, which are developer-written tests (also code) that test the classes and methods that are used. These are a form of “white-box testing” because they test the internals of the system and the various code paths that may be executed.



However, the philosophy of TDD applies equally well to story acceptance testing as it does to unit testing. This is called *acceptance test-driven development*, and whether it is adopted formally or informally, many teams write the story acceptance test first, before developing the code. The acceptance tests serve to record the decisions made in the conversation (card, conversation, confirmation) so that the team understands the specifics of the behavior the card represents. The code follows logically thereafter.

Proponents argue (correctly, we believe) that writing the acceptance test first is lean thinking that reduces waste and substantially increases the productivity of the team. This was illustrated best in the simple equation that Amir Kolsky, of NetObjectives, showed on a whiteboard. As shown in Figure 10–3, Amir wrote this:<sup>3</sup>

If  $R_t$  is 0, of course, then there is no savings. However, we all know that  $R_t$  isn't always zero, and since finalizing the test finalizes our understanding of the required behavior, why not write the test first, just to be sure?

Where:  
 $C_t$  is the time to write code  
 $T_t$  is the time to write test  
 $H_t$  is the time to hook test

Time to complete story if you write test first  
 $= T_t + C_t + H_t$

Time to complete the story if you don't  
 $= C_t + T_t + H_t + R_t$

Where  $R_t$  is the rework time necessary to pass the test once the test is understood and available.

**Figure 10–3** The simple math behind acceptance test-driven development

3. Personal interaction between Kolsky and Leffingwell

## ACCEPTANCE TEST TEMPLATE

At each iteration boundary or whenever a story is to be implemented, it comes as no surprise to the team that they need to create an acceptance test that further refines the details of a new story and defines the conditions of satisfaction that will tell the team when the story is ready for acceptance by the product owner. In addition, in the context of a team and a current iteration, the domain of the story is pretty well-established, and certain patterns of activities result, which can guide the team to the work necessary to get the story accepted into the baseline.

To assist in this process, it can be convenient to the team to have a checklist—a simple list of things to consider—to fill out, review, and discuss each time a new story appears. Crispin and Gregory [2009] provide an example of such a story acceptance testing checklist in their book. Based on our experience using this checklist, we provide an example from the case study in Table 10–1.

**Table 10–1** An Acceptance Testing Example from the Case Study

Story	
Story ID: US123	As a consumer, I always see current energy pricing reflected on my portal and on-premise devices so that I know that my energy usage costs are accurate and reflect any utility pricing changes.

Conditions of Satisfaction	
1. Verify the current pricing is always used and the calculated numbers are displayed correctly on the portal and other on-premise devices (see attachment for formats).	
2. Verify the pricing and the calculated numbers are updated correctly when the price changes.	
3. Verify the "current price" field itself is updated according the scheduled time.	
4. Verify the info/error messages when there is a fault in the pricing (see approved error messages attached).	

Modules Impacted	
Pricing RESTlet API	Impact: Amend protocol to allow pricing data.
In-home display	Impact: Refactor pricing schedule to support pricing programs to display on the in-home display.
Portal	Impact: Refactor pricing schedule to support pricing programs to display on the portal.

Documents Impacted	
User guide	Impact: Add new section on pricing.
Online help	Impact: Update online help to reflect pricing programs.
Release notes	Impact: Document defects in release notes.
Utility guide	Impact: Document pricing schedule changes.

Test Case Outline	
Test ID: <input checked="" type="checkbox"/> <b>Manual</b> <input type="checkbox"/> Automatic	Outline: 1. Check pricing: When there is no pricing info for a user: 2. Change of pricing: <ul style="list-style-type: none"><li>• When there is a pricing change in all allowed ways.</li><li>• Effective in the future.</li><li>• Effective in the past before the current pricing.</li><li>• Effective in the past but later than the current pricing.</li></ul> 3. Our current release does not support pricing change in the middle of a billing cycle. 4. Check the dashboard billing period consumption and the current bill to date.

Communications		
Internal	Involved parties: marketing, sales, product management.	Message: This is a new marketable feature.
External	Involved parties: utilities.	Message: This is a new feature to support new programs.

Since each team is in a different context, their templates will differ, but the simple act of creating a template as a reminder of all the things to think about benefits the team and increases the velocity with which they can further elaborate and acceptance test a new story.

## AUTOMATED ACCEPTANCE TESTING

Because acceptance tests run at a level above the code, there are a variety of approaches to executing these tests, including manual tests. However, manual tests pile up very quickly (the faster you go, the faster they grow), and eventually, the number of manual tests required to run a regression slows down the team and introduces delays in the value stream.

To avoid this problem, most teams know that they have to automate most of their acceptance tests. They use a variety of tools to do so, including database-driven tests, Web UI testing tools, and automated tools for record and playback.<sup>4</sup> However, many agile teams have discovered that some of these methods are labor-intensive and can be somewhat brittle and difficult to maintain because they often couple so tightly to the specific implementation.

A better approach is to take a higher level of abstraction that works directly against the business logic of the application and one that is not encumbered by the presentation layer or other implementation details.

### **Automated Acceptance Testing Example: The FIT Approach**

One such method is the Framework for Integrated Tests (FIT) method created by Ward Cunningham [Mugridge and Cunningham 2005]. This open source framework was designed to help with the automation of acceptance testing in a fast-moving agile context.

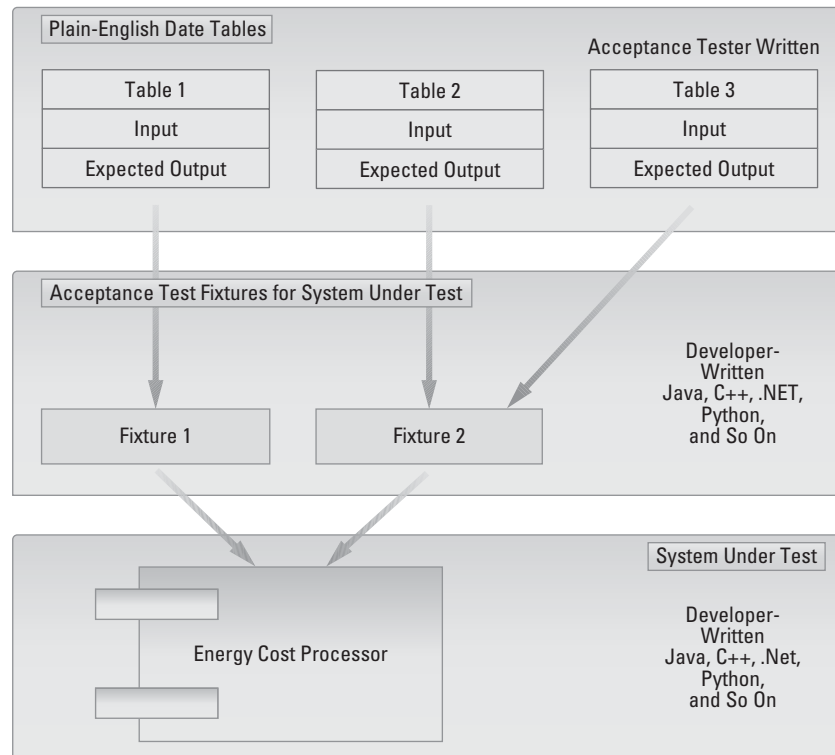
The FIT approach mirrors the unit testing approach in that the tests are created and run against the system under test, but they are not part of the system itself. FIT is a scriptable framework that supports tests being written in table form (any text tool will work) and saved for input as HTML. Therefore, these tests can be constructed in the business language (input and expected results) and can be written by developers, product owners, testers, or anyone on the team capable of building the necessary scripts.

Another open source component, FitNesse, is a wiki/web-based front end for creating text tables for FIT that also provides some test management capability. FIT uses data-driven tables for individual tests, coupled with fixtures or methods written by the developers to drive the system under test, as Figure 10–4 illustrates.

During the course of each iteration, new acceptance tests are developed and validated for each new story in the iteration, and these tests are then added to the regression test suite. These suites of acceptance tests can be run automatically against the system under test at any time to assure that the build is not broken by the new code.

---

4. Many agilists consider the record and playback model obsolete. Elizabeth Hendrickson notes three key reasons why traditional record and playback test automation solutions are not agile: “(1) The test-last workflow encouraged by such tools is all wrong for agile teams. (2) The unmaintainable scripts created with such tools become an impediment to change. (3) Such specialized tools create a need for Test Automation Specialists and thus foster silos.” [www.infoq.com/news/2008/05/testobsessed-agile-auto-testing](http://www.infoq.com/news/2008/05/testobsessed-agile-auto-testing)



**Figure 10–4** Acceptance test framework for the costing processor

As with unit tests, the FIT approach requires continuous involvement by the developers in establishing the acceptance test strategy and then writing the fixtures to support the test, illustrating yet again that the define/build/test team is the necessary structure for effective agile development. As always, the team's goal is to develop and automate tests within the course of the iteration in which the new functionality is introduced.

In some application domains, FIT does not provide an appropriately configurable framework. Sometimes, teams must build custom frameworks that mirror this approach but work in the technologies of their implementation.

In any case, whatever can't be automated eventually slows the team down, so continuous investments in testing automation infrastructure are routine items on a team's backlog, and like any other backlog item, they must be appropriately prioritized by the product owner to achieve sustainable high velocity.

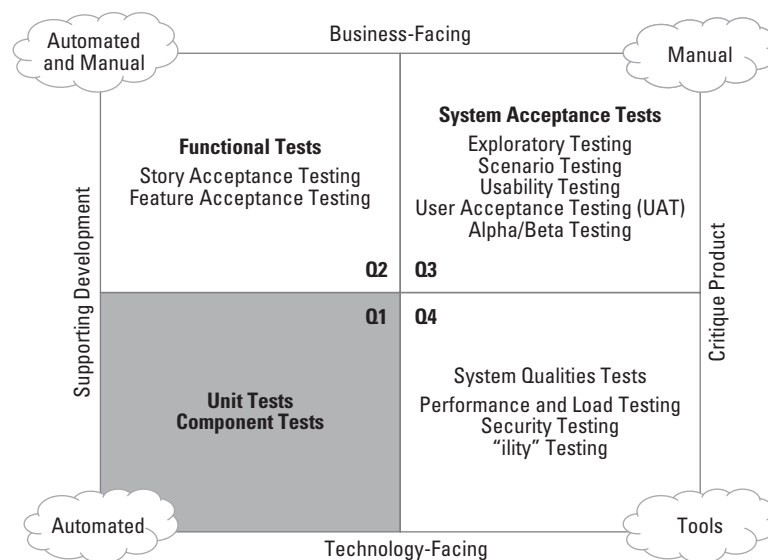
## UNIT AND COMPONENT TESTING

Before we leave this chapter, we must drill down one more level into testing and discuss quadrant 1 in the agile testing picture. In quadrant 1, we see *unit tests* and *component tests*, which are technology-facing tests as they are implemented and executed by the development team. Together with feature and story acceptance tests, they also complete the *support development* side of the agile testing picture, as Figure 10–5 illustrates.

### Unit Testing

Unit testing is the white-box testing whereby developers write test code to test the production code they developed for the system. In so doing, the understanding of the user story is further refined, and additional details about a user story can be found in the unit tests that accompany the code. For example, the presentation syntax and range of legal values for *current price field* can likely be found in the unit tests, rather than the acceptance tests, because otherwise the acceptance tests are long, are unwieldy, and cause attention to the wrong level of detail.

A comprehensive unit test strategy prevents QA and test personnel from spending most of their time finding and reporting on code-level bugs and allows the team to move its focus to more system-level testing challenges. Indeed, for many agile teams, the addition of a comprehensive unit test strategy is a key pivot point in their move



**Figure 10–5** Unit testing in quadrant 1

toward true agility—and one that delivers the “best bang for the buck” in determining overall system quality.

The history of agile unit testing closely follows the development of XP, because XP’s test-first practices drove developers to create low-level tests for their code prior to, or concurrent with, the development of the code itself. The open source community has built unit testing frameworks to cover most forms of testing, including Java, C, C++, XML, HTTP, and Python, so there are unit testing frameworks for most languages and coding constructs an agile developer is likely to encounter.

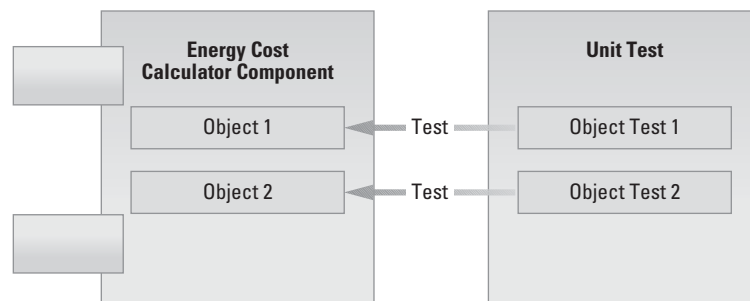
These frameworks provide a harness for the development and maintenance of unit tests and for automatically executing unit tests against the system under development. Unit testing the energy cost calculator might be a matter of writing unit tests against every object in the component, as Figure 10–6 illustrates.

The unit tests themselves are not part of the system under test and therefore do not affect the performance of the system at runtime.

#### Unit Testing in the Course of the Iteration

Because the unit tests are written before or concurrently with the code and because the unit testing frameworks include test execution automation, all unit testing can be accomplished within the iteration. Moreover, the unit test frameworks hold and manage the accumulated unit tests, so regression testing automation for unit tests is largely free for the team. Unit testing is a cornerstone practice of software agility, and any investments a team makes toward more comprehensive unit testing will be well rewarded in quality and productivity.

Figure 10–7 shows an example of a unit test for Tendril’s energy costing module that takes the consumer’s time-based pricing structure and runs a small sample that will exercise many of the costing module’s pathways. This test ensures that the costing



**Figure 10–6** Unit testing the energy cost calculator

```

@Test
public void testDailyCost_MultiplePrices() {
    List<ConsumptionValue> consumptionValues = new ArrayList<ConsumptionValue>();
    consumptionValues.add(new ConsumptionValue(TUESDAY_NOON, new Double(100)));
    consumptionValues.add(new ConsumptionValue(THURSDAY_NOON, new Double(200)));

    List<TemporalPrice> prices = new ArrayList<TemporalPrice>();
    prices.add(new TemporalPrice(new BigDecimal(".10"), SUNDAY, FIXED_PRICE));
    prices.add(new TemporalPrice(new BigDecimal(".25"), WEDNESDAY_NOON, FIXED_PRICE));

    DailyConsumptionHistory dailyConsumptionHistory =
        new DailyConsumptionHistory(new DayRange(SUNDAY, 7), consumptionValues, prices);

    DailyCost dailyCost = dailyConsumptionHistory.getDailyCost(new Day(WEDNESDAY_NOON));
    assertNotNull(dailyCost);

    /* First half of Wednesday is .10 / kWh, second half is .25 / kWh */
    /* 50 kWh burned that day = 25 * .10 + 25 * .25 = 2.50 + 6.25 = 8.75 */
    assertEquals(new BigDecimal("8.75"), dailyCost.getCost());
}

```

**Figure 10–7** A sample unit test from the case study

Thanks to Ben Hoyt of Tendril for this example.

algorithm effectively accommodates price changes that may occur at times other than standard day boundaries. Note that many more unit tests will be required before the costing module can be considered *done*. This is just one of many unit tests using the JUnit testing platform. This unit test example is one of many associated with calculating current costing, but it is automated and is an integral part of the regression testing package. The sample shows another useful feature in that there are comments embedded in the test to explain what the test results are supposed to be.

## Component Testing

In a like manner, component testing is used to test larger-scale components of the system. Many of these are present in various architectural layers, where they provide services needed by features or other components.

Testing tools and practices for implementing component tests vary according to the nature of the component. For example, unit testing frameworks can hold arbitrarily complex tests written in the framework language (Java, C, and so on), so many teams use their unit testing frameworks to build component tests. They may not even think of them differently.

Acceptance testing frameworks, especially those at the level of *http Unit* and *XML Unit*, are also employed. In other cases, developers may use testing tools or write fully custom tests in any language or environment that is most productive for them.



## SUMMARY

In this chapter, we described acceptance testing as an integral part of agile requirements management. If a requirement (feature or story) is not tested, unless it's simply work in process, which we indeed try to minimize, it doesn't really have any value to the team or to the user. With this discussion, we described the agile testing approach to assure that each new story works as intended, as it is implemented. This covers quadrant 2, functional testing of the agile testing matrix.

We also described the testing necessary to assure systematic quality from the perspective of quadrant 1, unit and component tests. Unit tests are the lowest level of tests written by the developer to assure that the actual code (methods, classes, and functions) works as intended. Component tests are higher-level tests that are written by the team to assure that the larger components of the system, which aggregate functionality along architectural boundaries, also work as intended. We also described how the team must endeavor to automate all the testing that is possible or else they will simply build a pile of manual regression tests that will eventually decrease velocity and slow down value delivery.

In Chapter 17, Nonfunctional Requirements, we'll describe acceptance testing practices associated with quadrant 4—system qualities tests.

*This page intentionally left blank*