Welcome to Patterns of Effective Test Setup!

This talk is about writing tests. My assertion, if you’ll pardon the pun, is that your tests suck.

Or, less insultingly, your tests suck up your time, suck up your employer’s money, suck the joy out of doing TDD, and just generally make your life more unpleasant than it needs to be. And they do that because you’re making a couple of very costly mistakes. You may not even recognize them as mistakes, but they’re causing you pain regardless.

It’s not your fault, though; lots of really smart people have written lots of really smart articles and books about how to write testable code and how to use TDD to drive the design of your code. But even if you were doing everything right, just like all those smart people said to do, I assert that you could *still* be making those mistakes. And that’s because many programmers tend to overlook “test setup” and test data as critical areas for innovation and improvement.

(click)

But that’s OK, I can help. And the reason that I can help is that I’ve spent a lot of time defining test setup patterns for my own team. I’ve made all of the mistakes you’re making right now. I’ve felt all the pain they create. And I have a solution that I think can save you that time, money, and frustration.

And the reason that I have this solution is that the project I manage has been under almost constant active development, week after week, sprint after sprint, for almost 8 years. Over that time the complexity of our code base, and the general size of our object model, has grown enormously. As a result, we really struggled for a while. The larger our object model got, the harder and more costly it was just to set up the test data for our tests, and the harder it was to maintain existing tests as we added new features. In fact, if we’d continued making those mistakes, instead of developing these new techniques, I don’t think we’d still be writing tests today. It would have become financially unbearable, and our software would have suffered greatly as a result.

My goal today is to open your eyes and give you a fresh perspective on your own tests. I want you to recognize the mistakes you’re making, I want you to be inspired to raise the bar and strive for clean setup code, and I want you to know how to get started when you get back to the office on Monday.

**(click for “Are you in the right place”)**

Before I get into the good stuff, I want to set some quick expectations about this session.

First, this is not a Testing 101 session. I assume that you’re familiar with at least the basics of writing tests.

Second, this isn’t about mocking or stubbing or how to write testable code. Those are really important topics, but lots of people smarter than me have written tons of words about that stuff already.

Third, I’m not trying to sell you any specific framework, library, or language. This is a talk about ideas and patterns. My slides will show you those ideas and patterns implemented using C# and NUnit, because that’s what I’m familiar with, but these concepts should be applicable in many different tech stacks.

Today, I’m going to focus entirely on improving the ways that you arrange your test data and prepare your system to execute a test. That’s it.

**Click for “Story time…”**

So why is test setup so important? The answer is because it makes up the majority of your test code.

Assuming that you’re not doing really bizarre in your tests, they likely all follow the same pattern: you do a bunch of stuff to get ready, then you call the one method or function that you’re testing and end with an assertion or two. Since most of the code you write deals with preparing your system for the test, then the primary drivers of how easily your tests are written, read, and maintained are the patterns that you follow in your setup code.

And if you’re not making deliberate choices to follow clean setup patterns, then you could be in for a world of hurt. To help illustrate that, I have a short story about why I’m here and where these ideas came from.

This story begins 8 years ago when I had just joined my current employer. We were just beginning our agile transformation and everyone was super excited about having “user stories” instead of “requirements” and “story points” instead of “estimates”. In the midst of all that agile euphoria, we decided to start writing automated tests for the new project we were starting.

Most of the team was new to testing and they didn’t have a good sense of what to test, or how many tests was “enough”. They felt that 100% coverage was unreasonable, but they were willing to commit to 70% coverage. At the time it seemed that having that warm, comforting safety blanket for 70% of our most important code seemed like a good start.

The project starts out great, everyone’s writing tests and shipping features and things are going pretty good. But a year or so later, after the code had started to get more complex and we’d started revisiting features to add new functionality, I began to feel that something was wrong with our tests. I’d really shotgunned the Agile kool-aid and was expecting this transformational impact from testing, but the tests just weren’t delivering that value.

As I investigated, I noticed that despite that code coverage commitment, many key objects weren’t being tested, or were barely tested. It turns out that as the code got more and more complex it got harder and harder to write the tests, so rather than testing the *most important* 70% of the code, developers were testing the *70% that was easiest to test*. And as you can imagine, that left a lot of important code uncovered.

One of the main reasons that tests were getting hard to write is that as the application got larger, it was requiring more and more effort just to describe the starting point for a given test. Complex business rules often required complex test data, and complex text data was time consuming to set up. So, people looked for ways to avoid testing the complex rules. But obviously, the complex rules is where test coverage was most important.

And when they couldn’t avoid the tests, they ended up writing some pretty gnarly setup code. In one particular case I needed to make a minor adjustment to a feature. The feature itself was complex, but the new change was relatively simple and I didn’t think it would take much time. Before writing any new code, however, I wanted to learn more about how the feature currently worked and I wanted to write a failing test. So I opened up the file containing the tests and my heart sank when I saw this:

**(click – code sample)**

and this….

(click)

And this.

That’s

(click) 29 string, integer and Boolean values being initialized,

(click) 7 different objects being created and

(click) 75 lines of code to understand.

(click) CRAP! And even worse, this is just the SHARED setup code for the test suite! Each individual test in the suite had more code like this, and each test depended on different portions of this mess.

It was clear that even though my change was simple, just modifying the existing tests would be difficult, let alone adding new ones. I went back to the team, increased my estimate, and spent way more time than should have been necessary implementing that change.

And this is not an isolated case! We have thousands of tests in our projects and we spend countless hours reading those tests and trying to make sense of stuff like this.

This sucks! But there is a better way.

**What’s on the agenda?**

Today we’re going to talk about what it means to have effective test setup patterns, we’re going to look at the mistakes you’re making today that reduce the effectiveness of your setup code, and then I’ll show you a number of patterns and techniques to do instead. We’ll finish by looking at ways for applying these same patterns and techniques to integration tests as well as unit tests.

**What is “setup”?**

But, let’s walk before we run. The first thing we need to do is answer these two questions. What do I mean by “test setup”, and how do we know if it’s being done effectively?

**(click for answer)**

By “test setup”, I mean anything that you do to create the baseline “input” for a test. This could mean creating objects in memory. It could mean putting data into a database. It could mean putting files on a file system. This could also mean setup code that’s shared between multiple tests or it could be setup code that’s unique to a specific test.

In general, I would include mocking and stubbing as “test setup”, but I’m focusing mostly on test data today.

**What is “effective setup”?**

When I say effective test setup, I’m referring to the art of writing clean, expressive setup that doesn’t suck.

**(click for version 2)**

Or, more precisely, I’m referring to coding patterns that increase the value that automated testing provides to your project

How do you identify effective test setup patterns? A project with effective patterns looks something like this:

**Signs of effective test setup**

First, tests are so easy to write, that you write a metric crapton of them. There’s something really enjoyable about getting into that TDD rhythm of red-green-refactor, but you can only do that if tests are painless to author. And if tests are painless to author, then either your code is simple or you’ve deliberately made them painless to author.

**Signs of effective setup #2**

Second, it’s a good sign if all of your tests are short and sweet. My rough rule of thumb is that the entire test should fit on the screen at one time. If your tests routinely require 2 or 3 screenfuls of code then my guess is that they are *not* easy to write, they’re probably not easy to *read,* and you probably aren’t writing a ton of them.

**Signs of effective setup #3**

Third, effective setup means that your tests don’t need a lot of refactoring or maintenance over time. Tests things are far less valuable if we’re constantly messing with them, and good setup habits can lead to more resilient tests.

**Signs of effective setup #4**

Lastly, a huge sign that you’re doing it right is that you can write integration tests that hit a real database or a real filesystem just as easily as you write in-memory unit tests. This is huge. This is the promised land. This is what I want to show you today.

Now, if I just described your project, then you’re probably in the wrong room because you’re already living in that promised land. However, I’m guessing many of you are here because your projects show signs of *ineffective* test setup.

**Signs of ineffective test setup**

The first sign of ineffective setup is that it’s too hard or frustrating or time consuming to do frequently. If testing isn’t fun, if you avoid writing tests because it sucks to do, maybe you’re doing it wrong.

**Signs of ineffective setup #2**

Another bad sign is if you curse in disgust every time you read or maintain an existing test. When I opened that original test I showed you a minute ago, I cursed like a sailor. I could tell instantly it was going to be a nightmare, and it was.

**Signs of ineffective setup #3**

A third sign of ineffective setup patterns is that tests frequently break, but it’s way easier to delete them then figure out how to fix them.

You ever do this? You’ve got a red dot on your screen, but you have no freakin’ clue what the test is doing, so you look over your shoulder, everyone’s at lunch, so you Ctrl-A, delete, Ctrl-S, git push commit. Boom. Fixed.

Yeah, that’s a sign that your tests need some work.

**Signs of ineffective setup #4**

And lastly, if your unit tests are painful to write or maintain, I’m guessing you don’t have a lot of integration tests. And if you don’t have integration tests, then you’re really missing out on some real-world feedback about how your system really works.

If any of these things sound familiar, then you’ve got some work to do. To help you with that, I’ve identified 4 mistakes that you might be making that make your tests so ineffective.

**Mistake #1**

The first mistake that people make when setting up tests is constructing all of their objects or test data by hand.

In almost every system there are tests that only care about a *portion* of an object. A test about an Order’s SHIPPING STATUS may not care about its line items, or a test about a Customer’s ADDRESS may not care about their name.

But it’s not always possible to create objects and specify ONLY what you care about. In C# for example the object’s constructor may require things that are necessary to the domain model, but don’t actually matter to that specific test.

**click**

For example, in this test, all I need is a shipped Order.

**click**

But apparently an Order needs a Customer…

**(click)**

And a Customer needs some Addresses…

**(click)**

And by the time I’ve satisfied the constructor, I’ve written a whole lot of code when I really only care about two things: the shipping status, and whether or not the order accepts new items.

**(click)**

All of these things that I created, but that don’t actually influence the assertion I’m making, are noise. Writing a test like this is painful, but it’s also painful to *read* these tests. You have to work hard to filter the signal from the noise so that you can understand it.

**(click)**

This mistake can also make your test code brittle.

What happens when the Order, Customer or Address constructors get modified? If you’ve ever made a simple change to your application, and then spent the next two hours cleaning up compilation errors and test failures in totally unrelated parts of the suite, then you’ve felt this pain.

**Mistake #2**

The second setup mistake is allowing your setup code to be noisy, complicated, or unclear.

One example of this is the test I showed you at the start. If your test requires 75 lines of dense object construction or mock setup, then you’re very likely doing something wrong. Those tests are a nightmare to write and even harder to read, as I discovered firsthand.

Another example of noisy and unclear setup code is when you create test data using explicit values, but those values don’t actually impact the outcome of the test.

**(click for example)**

Consider this example. The Customer and Product constructors both require certain values. And in addition, let’s assume that the test we’re going to run will execute code that throws an error if the customer ID or email address are left at their default values of 0 and null.

**(click for circled example)**

The problem is, when you set up this test, you’re forced to specify all those values even if they DO NOT MATTER to the test outcome.

When other programmers read that code, they have to spend time figuring out which values are significant and which are arbitrary. Are we testing a business rule that cares if the product is taxable or not?

And if this code is shared by multiple tests, it can be hard for someone to figure out which values they can modify to fit their needs and which values are significant to existing tests.

**Mistake #3: Inheritance**

The third mistake I see in setup code is using inheritance as a way of sharing logic between multiple tests.

I often find that there’s a certain amount of boilerplate setup that’s useful across multiple fixtures. For instance, you might create an Order, a Customer, and a few Line Items and link them all together in a meaningful way. This arrangement could be useful when testing *any* of those objects or any number of related business features, so naturally we’d want to make that setup logic reusable.

**(click for example)**

A quick and easy way of doing that would be to create a base class and do the setup there, like you see here. You’d expose the things you want to share as instance properties.

Every class that needs that shared data could derive from that base class and get access to the data.

There are two problems with this. First, inheritance is a very restrictive way of achieving reuse. In C# you can only have a single base class, and there’s just no good argument for requiring that your Customer tests and your Order tests derive from the same base.

**(click – the derived class)**

Secondly, as your system evolves over time, the needs of your tests might start to diverge. Maybe there’s one specific Customer test that needs to specify a distinct email address, or an Order test that requires that the Order not have any line items.

If the setup code is in the base class, you end up doing things like initializing data in the shared area, and then overriding parts of it in the body of each test.

**(click for circled)**

This just doesn’t scale over time. It’s confusing to read and maintain because you’ve taken this single logical thing, the instantiation of your test context, and you’ve split it across multiple files. And if anyone ever changes the shared data, your test might break until you add another override.

To properly reuse setup logic we need to get it out of a base class and into something more easily managed. I’ll show you what that looks like in a moment.

**Setup mistake #4**

The final mistake I see people making is writing integration tests that make assumptions about the state of external systems. This might mean assuming that a file exists in a specific location, or assuming that specific data will exist in the database.

When I first started doing integration tests, I basically created, in advance, a whole bunch of data that could be used for all of the tests in my system. I took a backup of that database, and I set up my tests to restore that database at the start of each test run.

In the tests, I had a massive file full of constants that were the primary keys of the various things I needed for different tests. I had the ID for a cancelled order, the ID for a successful order, the ID for an order with taxable items, the ID for an order using FedEx shipping, etc. Each integration test would then be assembled using those pre-built pieces.

That’s *madness*, for what I hope are obvious reasons.

**(click for example)**

What I’ve come to realize since then is that making *any* assumption about the state of external system is a huge mistake. Even something as simple as this can be a problem. This test needs a Customer, so it just grabs the first one in the database.

But what if you’re running against an empty database? Or what if this email service is designed to reject customers that have a known bad email address? Maybe on your database this pulls back a normal customer and the test passes, but on a coworker’s machine it pulls back someone in a flagged state and fails.

As a rule, every assumption makes your tests that much more brittle. My rule is that each integration test must set up everything that it needs and cannot rely on the database being in any specific state. I’ll show you how to do that in a minute.

**There is a better way!**

So how do we avoid those mistakes and write tests that *don’t* suck up all our time, money and energy?

I’ve identified 4 key practices that you can implement to avoid each of those mistakes.

**Key #1: Use helpers**

The first key to success, and the single most important thing you can do to improve your setup code, is to stop constructing test objects by hand. Instead, push object creation into some sort of helper method or object.

This gives you two benefits:

1. It often shortens your setup code, making it easier to write and read
2. It increases resiliency; if an object’s constructor changes, you potentially only need to update the helper method.

**(click – exploding req)**

Here’s an example of why this matters.

In my app, one of our core domain concepts is a thing called a “workflow”. There are very few things that a user can do that don’t involve a workflow in one way or another, which means that many of our requirements deal with workflows in different states.

As a result, many of our requirements look like this: “*When a workflow is <configured like this> then the system <should do that>*”.

**(click for explosing)**

In practice, though, actually creating a workflow in that state is complex; a workflow is composed of lots of smaller objects that work together, and they all have to be set up in a logically consistent way to represent real-world code paths and to avoid runtime errors.

When we have to create those objects by hand, we end up with the mess I showed you at the start. And since we deal with these objects all the time, anything we can do to make this type of setup faster pays huge dividends.

There are a couple of well-known patterns for handling object creation.

**Object Mother**

The first pattern that we tried is called Object Mother. The key idea behind this pattern is that you identify up front the different test data that you’ll need, and then you create static factory methods for each of those pre-defined states. For example, the “Order Mother” object might have a factory method for creating an order with an unpaid balance, or if you work with insurance, the “Policy Mother” object might create an insurance policy object with a specific combination of coverages.

Object Mother is a great way to get all of those noise values and objects out of your setup code, but it doesn’t really scale that well. As your software gets more complex you’ll need more and more pre-built objects in more and more pre-defined states. And as the number of pre-built objects and states grows it becomes harder to maintain them and harder for developers to choose between them.

(click)

Eventually, you end up with a mess like this. There’s one method for creating an order with different bill-to and ship-to addresses. There’s one for indicating that the credit card failed address verification. There’s one for specifying that the order was placed by a new customer.

There’s a ton of overlap here. What happens when someone needs an order that was placed by a new customer, and had failed the AVS checks, *and* had different bill-to and ship-to addresses? That exact scenario isn’t covered by any of these, so that developer would probably end up creating yet another method for their exact need. And that new method would probably have a lot of duplication when compared against the ones that already exist.

Object Mother is a really easy pattern to implement if you only need a couple of course-grained pre-built objects. We needed a lot more control over our test data, so we quickly outgrew this pattern.

**(click for Data Builder)**

The next thing we tried was a pattern called Data Builder.

Rather than a factory that returns pre-built objects, Data Builder lets you create customized objects in the body of each test. It’s common for this to be accomplished via a Fluent API that exposes the things that can be customized.

The general structure in this pattern is that initialize the builder itself and then start calling methods to customize various parts of the object. Those methods are chained together and at the very end you call a Build() method which returns your fully built object.

These can be very simple, or they can get pretty complex as you see here where we’re creating both an Order and Customer with customized properties.

The main benefit of the Builder pattern is flexibility because it lets you can create the precise data that you need for each test, and that makes it a much better fit for larger or more complex applications.

**(click for fluent API)**

However, I’m really not a big fan of the Fluent API. It’s verbose and adds a lot of noise, as you see here. The green arrows are pointing to the significant data that I’m creating and the red circles are basically the “noise” that we get from the fluent API.

All this noise code means that the setup code is harder to write, read, and maintain. And on top of that, actually implementing the Fluent API is tedious. It requires a lot of boilerplate code that in my experience just isn’t worth the hassle.

**(click for Test Helper)**

Eventually, we created a hybrid of these two patterns that combines the static factory class of Object Mother with the customizable nature of a Data Builder, minus the Fluent API.

We call this the Test Helper pattern. It’s a terrible name, but it’s been a really useful pattern for us and I think it’s highly applicable not only in C# but also in JavaScript or Ruby.

**(click)**

The first step in implementing a Test Helper is to create a static factory class, one for each type of object you want to build. I generally name them like this: *FooHelper* creates a *Foo*.

That factory gets a single static method called “Create”, which allows the caller to specify method arguments to customize the object. In this case, I’m using the “named arguments” feature of C# to specify which arguments I’m providing. If your language doesn’t support this feature then you might need to use method overloading to achieve something similar.

**Click – meaningful defaults**

The helper’s job is to build the object, using meaningful defaults for whatever isn’t passed as arguments. In this case, I’ve opted to assign a specific status and name by default.

I generally declare *all* arguments with a null default and then use the null coalescing operator when constructing the object. You could assign the default value right here in the signature, but sometimes it’s handy to tell if the caller provider a value or not.

**(click – delegate**)

If the object you’re building is composed of other objects, then expose them as optional arguments as well. If the caller leaves it null, then delegate to the relevant helper to create those objects.

This lets your test code declare *only* what it actually needs, and then the helpers fill in all the gaps. And each helper only needs to deal with one type of object. This keeps both your tests and your helper code clean and tidy.

**(click – accidental equality)**

When assigning default values, there’s one potential gotcha you should be aware of. I call this the problem of “unexpected equality”.

For example, let’s say you create two different Customers from the helper, you pass them into some method that performs a comparison, and that method returns the ID of the customer matching some business rule. Your test asserts that the return value is equal to the customer that meets that requirement.

The expectation here is that the test should *fail* if the code returns the ID of the *other* customer you created. But if the CustomerHelper creates Customers with a static, hardcoded ID value, then both Customers will be created with the same ID. And in that case, the test will pass even if the code is broken. This is what I mean by “unexpected equality”; if I create two separate objects, I don’t expect them to be considered equal.

So, by default, I prefer to make all significant values unique. This applies to IDs, names, email addresses, etc. I want programmers to intentionally be explicit if they want things to be equal.

**(click – short guid)**

To assign unique strings, I use a class called ShortGuid. This is basically a shorter, URL-friendly, base64-encoded GUID. It’s more compact than a standard GUID, which can be helpful in some cases, but it’s still guaranteed to be unique. You can get this code off the web.

Whenever I’m creating a name or a title or something, I use a ShortGuid as the default. It guarantees that no two objects I create will share the same value, unless I explicitly set them up that way.

**(click – id sequencer)**

To assign unique integers I created a static class called the IdSequencer. It basically starts a counter and hands out a unique value each time I call “next”. Any time I have a helper that creates something with an ID property, I expose the ID as an argument and then I default it using the sequencer.

You *could* assign a random number instead, but this sequencer comes in really handy on integration tests. You’ll see that in a few minutes.

**(click – Test Helpers vs Libraries)**

People often ask me why I write my own Test Helper classes instead of using a 3rd party object construction library. In .NET there are libraries that take a generic type argument and then automagically create an instance of that type, populating it with test data. This particular example is from a library called AutoFixture, and it looks super easy. Why not do this?

The simple answer is that no library that you download can make better decisions about your test data than you can. Early on, when your objects are simple and you don’t have lots of special cases in your system then sure, this might work OK. But as things get complex, you’re going to want control over how your default values and properties get set up, and you don’t get that with a library.

And in addition, these libraries tend to be noisier than custom code.

**(click – Test Helpers vs Libraries, #2)**

Here’s what AutoFixture looks like if you want to specify a value for a property. Sure is a lot simpler to read and write the Test Helper version.

Basically, I look at it like this: when your app is small and simple, then the libraries might work. But if your app is small and simple, then introducing my Test Helper pattern is super easy too. And since Test Helper is designed to deal with complexity, you’ll start out with a solid foundation that will scale right along with your app code. And if your app is already really complex, then these libraries aren’t going to save you as much time as you might think.

**Key Practice #2: Tell a story**

The second key to success is to make sure that your test setup code, and your test data, tell a story.

The reason is that, at the core of their essence, tests are valuable because they help us understand our software. And in order to fully deliver that value, they have to effectively convey information when they are read. You could write the most bassackwards and incoherent tests possible and the computer could still figure out if the assertions hold true. But that won’t help your poor coworker who opens that file a month later and needs to make a change. CPU cycles are cheap, your coworker’s time is not. Investing in clear, concise setup code is key.

When it comes to telling a story, there are a few basic things to keep in mind.

(click)

First, names matter. If your setup code is a story, then your variable names and values are the actors in that story. Use them to drive home the *point* of a given test.

In this example, the actual values in use are irrelevant; they are just two arbitrary values that need to be different. By giving those values names, the assertion now has a little extra clarity and clearly supports the purpose of the test.

(click)

Or if your test revolves around the fact that a customer is disabled, don’t just call it “customer” or “c”; call it “disabledCustomer”.

(click)

If your test contains multiple objects of the same type, differentiate them! It sounds obvious, but I still see smart, experienced developers writing tests with variables called “customer1” and “customer2”. That is a completely unnecessary mistake that reduces the clarity of your setup code.

In this case, I’m testing some search logic. To write a complete test, I need to assert both that the code DOES return a match and DOES NOT return something that doesn’t match. Naming the objects in this way provides way more meaning than “customer1” and “customer2”. And if you use consistent names across your tests, your coworkers will start to recognize these patterns. If I see the word “distractor” in any of our tests, I know *exactly* what it’s there for.

(click)

Here’s one final example of how good names can convey meaning. This test is asserting that a specific feature returns data ordered in a specific way.

If I create the test data in the same order in which I expect them to come back out, then it’s possible that the sorting code isn’t do anything at all, and the test is passing by coincidence. A better test would be to create the data in a *different* sequence than it should come out, because then the test will only pass if it’s actually applying some sort logic.

Here’s one way to write that test. This isn’t terrible, but it’s not telling as clear of a story as it could.

(click)

Here’s that same test, but where I assign those values names that describe their purpose. And as a result, I’m able to write an assertion that far more explicitly captures my intent behind the test.

On the sorts of tests that fit onto these slides, some of these techniques may not seem that useful. But if you get in the habit of doing this, then you’ll start to see a real difference in the readability of your real-world tests, especially as your setup logic gets larger and more complex.

**(click – clean, concise API)**

Another way to tell a clear story is to use your helper API to write fewer lines of code, but at a higher level of abstraction. Here’s an example where I’m using a test helper to set up an Order that has two payments associated with it. Creating this order is a single logical concept, but it requires multiple physical statements to accomplish.

Also, in order to make sure that this object is internally consistent, I might have to make sure that the order subtotal

**(click)**

By taking a more declarative approach I reduce the amount of code that I have to write and I can clearly describe the data I need in a single statement.

In addition, note that I no longer need to manually specify the order subtotal; the helper can easily infer that value from the sum of the two payment amounts and wire everything up for me.

And once this piece of code is written, it’s available for reuse by other tests that have similar needs.

**Key #3: Scenarios**

Test Helpers are great at returning single objects. But what if you need of keep track of multiple objects AND their relationships?

For example, let’s say you have an ecommerce site, and one of your business rules is that all orders of heavy equipment, from new customers, with a different bill-to and ship-to address, must go through a verification process to prevent fraud and expensive shipping mistakes. To write that test, you’d have to create a customer with no previous orders, assign different bill-to and ship-to addresses, create an order containing a heavy equipment item, and attach the customer to the order.

Wiring all that stuff up by hand is tedious and works against our goal of being able to easily and concisely describe the context for a test.

And if you have multiple tests that need minor variations on this setup, you end up with a lot of copying and pasting.

(click)

In these situations I use a pattern that we call a Scenario. This is essentially a façade that wraps the coordination of multiple Test Helpers towards a common goal and makes your setup code cleaner and more readable.

You’ll notice that while a Test Helper is a static factory, a Scenario is something that you instantiate. You can still customize the result but you do it with constructor arguments and not method arguments.

The reason for this difference is that a Test Helper returns one of our core domain objects, and we don’t want to litter our app code with constructors that exist only for testing. The factory pattern works great to isolate the Test Helper logic from the core objects.

For scenarios, though, we’re actually creating *multiple* objects, and we need a handy way to keep track of all of those objects. If we implement the Scenarios as brand new classes, then we can use instance properties of those classes to expose pointers to the objects the tests will care about. You could still use static factory methods if you wanted to, but it saves a little code to just use the constructor instead.

(click)

This is what the Scenario itself looks like. In general, everything that the Scenario creates that a test might need to easily get a reference to is exposed as instance properties.

(click)

Here’s another example of a Scenario. In this case, we’re creating multiple orders for a single customer, each with different characteristics. The Scenario exposes each order as a distinct property so that the test code will be very clear in terms of its intent.

(click – drawbacks)

I recommend the Test Helper pattern without reservation. There’s literally no good argument not to. But there are a few drawbacks to the Scenario pattern that you should consider.

First, you’ll need to figure out for yourself the correct balance between specialization and customization. If you create a large number of Scenarios, each highly specialized for a specific use case, then you’ll generally find that each individual is scenario easy to maintain over time, but you’ll end up creating a lot of them with chunks of duplicate logic. If you create a smaller number of general purpose Scenarios that can be highly customized via arguments then you’ll create fewer classes, but it may be hard to refactor a Scenario because it might be used in lots of different ways by lots of different tests.

That balance is hard to predict in advance. My rule of thumb is that if I’m only creating two or three objects in a test, I’ll usually just call the Test Helpers directly. But if I have more than that, AND I’m reusing that logic in more than two or three tests, I’ll extract out a Scenario.

Also, remember that Scenarios are helpful because they are facades that encapsulate and abstract the coordination of multiple objects at the same time. If you need a lot of Scenarios, that might be a code smell indicating that those facades would be more useful in the core app itself. For example, if we go back to that e-commerce scenario I just showed you, you could argue that a better design might be to create an object or service that encapsulates all of the information about a given Order, and then base your business rules on that abstraction instead. If you go that route then you may need fewer Scenarios in your unit tests, but they are still really handy when it comes to integration tests.

**Key #4: Integration tests**

Until now we’ve been talking about in-memory objects only. The fourth key to effective setup is to leverage those in-memory helpers in your integration tests as well.

This is a key practice because it addresses mistake #4 that we talked about. If you have a rich, customizable set of helpers that create data, then your integration tests no longer need to rely on pre-existing data. Instead, they can create exactly what they need in exactly the shape they need it.

If, like me, you have a standard object oriented web app, then being able to use the same set of helpers for in-memory tests as I do for integration tests is huge. If you’re doing a lot of functional programming, or if you tend to abstract things away into services, then you may not need complex data objects for unit tests. But regardless of your app’s architecture, however, eventually you’re going to want to test your data access code or do some other system level integration testing. And when that time comes, having a set of data creation helpers that can populate a database is really useful.

Unfortunately, it’s easier said than done.

**(click – Foreign Keys)**

First, you have to deal with foreign keys. Your app may not care if you create a Line Item by itself, but you can’t save the Line Item to the database without an Order. And maybe you can’t create an Order without a customer. It’s the same issue we had with constructor dependencies, but in the database.

This means you have to new up the entire object graph and then save objects to the database in the correct sequence.

**(click – Primary keys**)

And if your database assigns primary keys, then after you save all those objects you have to update their ID values with the newly assigned key.

**(click – column constraints**)

You also have to make sure that your helpers are creating data that can be saved. Some columns might have constraints that reject NULL or other values, and those constraints may not be duplicated in the domain model.

**(click – junk data)**

Lastly, you’ll want to clean up that test data when the test run is over. I run my automated tests against the same database I use for manual testing. I don’t want that database filled up with junk data because it can impact performance, waste disk space, and it’s just ugly.

These things make integration tests difficult, but we can handle them with a few extra additions to the Test Helper pattern.

**(click)**

The first thing is to add a *Save()* method to your Test Helpers classes, and pass your database connection into it.

We use NHibernate and this ISession thing is basically the database gateway object. If you use Entity Framework then you might pass the db context here, or you might pass a raw ADO.NET connection or whatever other object that you need to talk to the database.

**(click – details of the Save #1)**

Here’s what the Save method itself looks like.

The first thing it does is deal with the foreign key constraints by delegating to other helpers to save its references. In this case, we can’t save an Order unless it references an existing Customer ID. By delegating to the Customer helper to create that record we keep each individual helper clean and focused on a single type.

**(click – details of the Save #2)**

Next, the Save method deals with primary key values.

Remember that to avoid “unexpected equality”, each object that we create is assigned a non-zero value. But many ORMs use the ID property to determine if they should issue an INSERT or UPDATE query, and if the ORM sees a non-zero ID it will issue an UPDATE statement, and not an INSERT.

This is where that IdSequencer object comes in handy. It knows which values it’s handed out, so before we save our object we do a check to see if the ID currently in use was assigned by the sequencer. If it was then we reset it to 0, and cause an insert. If it wasn’t, then that means we’re dealing with an object that already exists in the database and we do nothing, and cause an update.

Note that we don’t need to reset ALL values that were assigned by the IdSequencer, only entity IDs. Only properties that map to primary keys.

**(click – details of the Save #3)**

Finally, the helper calls out to the ORM to insert or update the database. If you’re not using an ORM then the exact details might change, but the general pattern should hold up.

If you follow this pattern then you can basically write a bunch of unit tests with in-memory data, copy and paste the setup code from one of them into an integration test, add a few calls to this Save method, and you’re done. It’s pretty sweet when it comes together.

But how do prevent this test data from lingering in the database when the test run is over?

One possibility is to reset the database to a known state at the start of each test run. This works, but I don’t recommend it. One reason is that it’s a massive pain to maintain that baseline backup every time the schema changes or new data is added. Another reason is that I use the same database for unit tests as I do for manual testing. It really sucks to spend a bunch of time crafting data for a manual test and then lose it because you accidently ran an integration test that wiped the slate clean.

**(click - Rollback)**

Another option is to start a database transaction when each test starts, and then roll that transaction back when the test is over.

Years ago I wrote a custom NUnit attribute called Rollback that does this for us. Any test that has this attribute is automatically executed inside of a transaction that is discarded when the test finishes. The implementation for this is on my GitHub, but at this point it’s like seven years old. There are probably newer and better ways of doing it now, but you’re welcome to copy my approach if you’re using NUnit.

**(Click – order to chaos)**

As an example of how powerful these techniques can be, let’s first revisit that nasty chunk of setup code I showed at the start…

(click)

(click)

All three screenfulls of it.

(click)

Here’s that same chunk of code, cleaned up and rewritten using Test Helpers. Most of the values and objects being created were irrelevant dependencies that didn’t impact the outcomes. After pushing all of that stuff into helpers, this is all that’s left and it’s way, way more readable.

I’ve said it before and I’ll say it again; the single most important thing you can do is build a good helper library and stop creating data by hand.

**How to get started?**

To wrap up, I have some quick suggestions for how to get started with these patterns in your own code.

First, start by creating helpers for your simple objects first, the ones that don’t have lots of dependencies or child data. Then move up to more complex objects, delegating to the simple helpers as needed. If you try and start with that huge, massive, ancient beast that lives at the heart of your legacy system, it’s gonna hurt.

Second, continually refactor your helpers as needed. Remember that test code *is* “real code”; keep it clean and tidy just like you would anything else.

Lastly, the sooner you start implementing these patterns, the sooner you’ll notice the payoff. There *is* an investment to add these to a legacy system, but the promised land of clean, simple unit AND integration tests is totally worth it.

**Closing**

And that brings us to the end of this session. Here are the 4 keys to effective test setup as well as links to these slides on Github, my website, and my twitter account. If you have any questions or comments I think we have a few minutes right now, or please feel free to seek me out online.

THANK YOU!