5. Testing

Bradley Dilger
Purdue University

Definition and Background

For many technical communicators and designers, testing is the most advanced step in the design thinking process, following research, *prototyping*, and development (Pope-Ruark et al., 2019). The basic framework is simple: Ask users of a document, product, and/or system to use it, observe and guide their interaction, record how well they achieve design goals, and use that information to guide future design work. Testing, then, is one of the strongly data-driven components of design thinking, complementing user research and prototyping in empowering designers to make design decisions based on real world use. By providing direct feedback about users' engagement, testing can be the engine of *iteration* that should be at the heart of design thinking.

Testing is often strongly identified with usability testing, and particularly the *usability* of websites, given the widespread influence of Jakob Nielsen (1997, 1999), Steve Krug (2009, 2013), and Carol Barnum (2001, 2020). As Barnum (2020) observes, the ISO definition of usability is the benchmark: "The extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (p. 11). Most designers have moved away from a rigid focus on usability to broader thinking about user experience, and testing has followed suit.

The researchers noted above have popularized a streamlined "discount" approach to testing which limits the complexity and number of test sessions in order to simplify testing and hopefully make it more common. While acknowledging that larger quantitative studies are sometimes necessary or desirable, Nielsen, Krug, and Barnum explicitly (and repeatedly, and sometimes stridently) argue that expensive, laboratory-like test conditions are not necessary for designers to achieve meaningful results. "Zero users give zero insights," writes Nielsen (2000). "As soon as you collect data from a single test user, your insights shoot up and you have already learned almost a third of all there is to know about the usability of the design." For these researchers, the limitations of small-scale testing using informal methods are far exceeded by the benefits of working directly with participants to learn about a product's successes and pinch points firsthand.

While testing websites for usability often dominates conversations about testing on design and user experience websites, or even in textbooks (e.g., Markel & Selber, 2018), anything built with a design process can benefit from testing—instructions, smartphone apps, promotional materials, or election ballots (see

Rachael Sullivan's **design ethics** example, or Jarrett & Redish, 2020). Testing can measure accessibility, persuasion, error tolerance, and more, and other research methods can be integrated into the process. For example, short posttest surveys can measure how well users remember critical information. Focus groups formed from multiple testers can shape the next steps for design teams. And testing can come at any stage of the design process, engaging wireframes, prototypes, or designs released to the public. See Digital.gov for widely used resources that describe how to conduct tests and demonstrate how testing can be integrated into design processes: for example, consent forms, materials for conducting practice tests, and case studies from experienced practitioners—many once part of the now-archived Usability.gov.

■ Design Application

In a typical test as described by Nielsen, Krug, and/or Barnum, two facilitators work with a participant in a comfortable setting, perhaps even a coffee shop or employee break room. One facilitator guides the participant through a pre-prepared list of tasks; the other takes notes and/or operates recording equipment such as a camera (for testing paper prototypes, products, or documents) or screencasting software (for testing apps, web pages, or software). For example, the facilitator may ask a participant to envision themselves in the following scenario and complete a set of tasks:

> Sample scenario:
>
> We are testing a mobile app for a pizza restaurant. Please use the app to order two pizzas for a party. Have the pizzas delivered. Please use the address and credit card number on this note card.
>
> Tasks:
>
> 1. Start a new order and add a plain cheese pizza.
> 2. Add a second pizza with black olives on one side and green olives on the other.
> 3. Enter your address and save it as the "Home" address for future orders.

When completing the tasks, participants are often asked to use the think-aloud protocol, a technique where they explain their actions and intentions "out loud" as they go, offering a richer data stream to facilitators. Generally, test facilitators do not ask questions during the session, but might do so in a short debriefing afterward. Participants are typically offered small incentives such as a gift card. The process is repeated for a small number of participants, usually five, and then a brief report is written to share with clients and designers if they are not directly involved or present to observe the test process.

As noted above, advocates of the "discount" approach to testing emphasize the goal is not generalizable research but providing designers with data-driven guidance for iteration. For this reason, testing best practices emphasize testing with a small number of participants and focus less on identifying test participants representative of user populations. Krug (2010) suggests other forms of user research can adjust for any inaccuracies arising from "non-representative" users. Indeed, he lowers the oft-quoted number of five users per test session—established by Robert Virzi (1992) and confirmed by Nielsen (2000)—to three, arguing that increased iteration is more likely to find ways to achieve designers' goals. "Recruiting loosely," as Krug describes it (2010, p. 42), is also an opportunity to ensure designers are not excluding historically marginalized populations when identifying test participants.

▮ Pedagogical Integration

Testing is a common exercise in technical communication courses (Summers & Watt, 2015), and learning to use testing for diverse purposes has many potential future uses. Direct contact with end users can help draw attention to other parts of the design process, such as *empathizing* with audiences, and the hands-on nature of testing makes it ideal for active learning. Particularly in *usability* studies or user experience research courses, students may be asked to devise methods and conduct product testing (Zhou, 2014). As with any pedagogical activities in technical communication, instructors should remind students of the critical and ethical dimensions of design. For testing, ethics is especially important, given that the sources mentioned above and testing resources commonly found using web searches focus heavily on web usability, where utility, expediency, and functionality are prioritized, sometimes carelessly.

For an introductory technical communication course, testing can be situated as a module or course unit, or can be integrated into larger projects to encourage iteration and data-driven thinking. Either way, practicing testing in class can scaffold learning by providing a ready source of test participants and opportunities for mutual assistance—students can rotate between facilitator, assistant, and participant roles, gaining perspective about the challenges of each. Available literature on remote testing (e.g., Moran & Pernice, 2020) facilitates its application in online or hybrid courses.

Though the widespread adoption of "discount" methods (Nielsen, 1997) has reduced barriers to testing, it remains labor-intensive, and technical communication instructors must allow adequate time for students to plan tests, analyze data, and draft reports. Students should seek feedback from their instructors, peers, or collaborators at each of these stages. Though testing takes a lot of work, the insights it provides are almost always worth it. Carefully planned testing not only makes better documents and products, but ensures they work for all audiences and keeps the needs of many different users in mind. Joseph Bartolotta et al.

(2018) and sources mentioned above (Summers & Watt, 2015; Zhou, 2014) offer more guidance for technical communication instructors seeking to meaningfully integrate testing into their curricula—and to continue the necessary work of broadening the focus of testing from web usability.

■ References and Recommended Readings

Barnum, C. M. (2020). *Usability testing essentials: Ready, set . . . test!* (2nd ed.). Morgan Kaufmann.

Barnum, C. M. (2001). *Usability testing and research*. Longman.

Bartolotta, J., Newmark, J., & Bourelle, T. (2018). Engaging with online design: Undergraduate user-participants and the practice-level struggles of usability learning. *Communication Design Quarterly*, *5*(3), 63-72.

Jarrett, C., & Redish, G. (2020, May 4). How to test the usability of documents. *UXmatters*. https://www.uxmatters.com/mt/archives/2020/05/how-to-test-the-usability-of-documents.php

Krug, S. (2010). *Rocket surgery made easy: The do-it-yourself guide to finding and fixing usability problems*. New Riders.

Krug, S. (2013). *Don't make me think, revisited: A common sense approach to web usability* (3rd ed.). New Riders.

Markel, M., & Selber, S. (2018). *Technical communication* (12th ed.). Bedford/St. Martin's.

Moran, K., & Pernice, K. (2020). *Remote moderated usability tests: How to do them*. Nielsen Norman Group. https://www.nngroup.com/articles/moderated-remote-usability-test/

Nielsen, J. (1993). *Usability engineering.* Academic Press.

Nielsen, J. (1997, January 1). *Discount usability for the web*. Nielsen Norman Group. https://www.nngroup.com/articles/web-discount-usability/

Nielsen, J. (1999). *Designing web usability*. New Riders.

Nielsen, J. (2000, March 18). *Why you only need to test with 5 users*. Nielsen Norman Group. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Pope-Ruark, R., Moses, J., & Tham, J. (2019). Iterating the literature: An early annotated bibliography of design-thinking resources. *Journal of Business and Technical Communication*, *33*(4), 456-465.

Summers, S., & Watt, A. (2015). Quick and dirty usability testing in the technical communication classroom. In *2015 IEEE International Professional Communication Conference (IPCC)* (pp. 1-4). IEEE.

Virzi, R. A. (1992). Redefining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, *34*, 457-468.

Zhou, Q. (2014). "That usability course": What technical communication programs get wrong about usability and how to fix it. *Communication Design Quarterly*, *2*(3), 25-27.