Don’t forget about human factors: Lessons learned from COVID-19 point-of-care testing

Sarah Farmer, Georgia Institute of Technology
Victoria Razin, Georgia Institute of Technology
Amanda Foster Peagler, Georgia Institute of Technology
Samantha Strickler, Emory University
W. Bradley Fain, Georgia Institute of Technology
Gregory L. Damhorst, Atlanta Center for Microsystems-Engineered Point-of-Care Technologies
Russell Kempker, Emory University
Nira R. Pollock, Atlanta Center for Microsystems-Engineered Point-of-Care Technologies
Oliver Brand, Atlanta Center for Microsystems-Engineered Point-of-Care Technologies
Brooke Seitter, Atlanta Center for Microsystems-Engineered Point-of-Care Technologies

Only first 10 authors above; see publication for full author list.

Journal Title: Cell Reports
Volume: Volume 2, Number 5
Publisher: Elsevier | 2022-05-23
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1016/j.crmeth.2022.100222
Permanent URL: https://pid.emory.edu/ark:/25593/vxrs5

Final published version: http://dx.doi.org/10.1016/j.crmeth.2022.100222

Copyright information:
© 2022 The Author(s)
This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/rdf).
Perspective

Don’t forget about human factors: Lessons learned from COVID-19 point-of-care testing

Sarah Farmer,1,2,3,16 Victoria Razin,2,3,16 Amanda Foster Peagler,1,2,3,16 Samantha Strickler,3,4,16 W. Bradley Fain,1,2,3 Gregory L. Damhorst,3,5 Russell R. Kemper,3,5 Nira R. Pollock,3,5,6 Oliver Brand,3,7,8 Brooke Settier,3,8 Stacy S. Heilman,3,16 Eric J. Nehl,3,16 Joshua M. Levy,3,16 David S. Gottfried,3,7 Greg S. Martin,3,7,8 Morgan Greenleaf,3,5 David N. Ku,3,12 Jesse J. Waggoner,3,16,13 Elizabeth Iffrig,3,16 Robert G. Mannino,3,16,13 Yun F. Wang,3,16,13 Julie Sullivan,3,9 Paulina A. Rebolledo,3,5,10 Viviana Claverı´ a,3,12 John D. Roback,3,14 MacArthur Benoit,3,8 Cheryl Stone,3,8 Annette Esper,3,16 Filipp Frank,3,15 and Wilbur A. Lam3,16

1Center for Advanced Communications Policy, Georgia Institute of Technology, Atlanta, GA, USA
2Georgia Tech Research Institute, Georgia Institute of Technology, Atlanta, GA, USA
3The Atlanta Center for Microsystems-Engineered Point-of-Care Technologies, Atlanta, GA, USA
4Department of Emergency Medicine and Department of Anesthesia, Division of Critical Care, Emory University School of Medicine, Atlanta, GA, USA
5Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA
6Department of Laboratory Medicine, Boston Children’s Hospital, Boston, MA, USA
7Institute for Electronics and Nanotechnology, Georgia Institute of Technology, Atlanta, GA, USA
8Children’s Healthcare of Atlanta, Atlanta, GA, USA
9Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA
10Rollins School of Public Health, Emory University, Atlanta, GA, USA
11Department of Otolaryngology-Head and Neck Surgery, Emory University School of Medicine, Atlanta, GA, USA
12Department of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA
13Wallace H. Coulter Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, USA
14Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, GA, USA
15Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, USA
16These authors contributed equally

*Correspondence: wilbur.lam@emory.edu

https://doi.org/10.1016/j.crmeth.2022.100222

SUMMARY

During the COVID-19 pandemic, the development of point-of-care (POC) diagnostic testing accelerated in an unparalleled fashion. As a result, there has been an increased need for accurate, robust, and easy-to-use POC technologies. While stakeholders often express the desire for POC technologies that are “as simple as digital pregnancy tests,” there is little discussion of what this means in regards to device design, development, and assessment. The design of POC technologies and systems should take into account the capabilities and limitations of the users and their environments. Such “human factors” are important tenets that can help technology developers create POC technologies that are effective for end-users in a multitude of settings. Here, we review the core principles of human factors and discuss lessons learned during the evaluation process of SARS-CoV-2 POC testing.

INTRODUCTION

Over the last several decades, point-of-care (POC) technologies have facilitated decentralization of diagnostic testing away from hospitals to medical offices (Murray et al., 2004; Price, 2001; Rayment et al., 2012), pharmacies (Gubbins et al., 2014), and even the home (Murray et al., 2004; Kennedy et al., 2017; Robinson, 2016). The COVID-19 pandemic has accelerated this evolution, moving POC technologies into the spotlight, as the medical community and general public have come to recognize the value of diagnostic testing in non-hospital settings (Kim et al., 2021; Parcell et al., 2020; Tromberg et al., 2020). As this shift has occurred, the rapid reporting of results provided by POC technologies has enabled faster clinical decision making and implementation of public health measures (Chowell et al., 2021). As such, POC diagnosis has flourished during the COVID-19 pandemic.

In contrast with POC testing, standard practice in diagnostic testing involves collection of patient samples by trained staff and processing in centralized or specialized clinical laboratories by medical technologists. This laboratory environment is standardized and regulated by government and professional agencies to maximize quality and minimize diagnostic error. Testing outside of a centralized clinical laboratory presents inherent challenges. In addition to potential loss of sensitivity as compared to standardized technologies, as POC technologies enable use in less...
controlled environments by users with less training, quality is less regulated and diagnostic errors may increase. Therefore, in order for POC diagnostics to be accurate in real-world conditions, they must be robust and easy to use.

This need for robustness yet simplicity is the premise expressed in being “as simple as digital pregnancy tests,” but what this means in practice can be unclear for technology developers and clinicians. To clarify this, we turn to the field of human factors, as defined by the FDA as the integration of “the knowledge about human behavior, abilities, limitations, and other characteristics of medical device users to the design of medical devices ... to enhance and demonstrate safe and effective use” (FDA, 2016). Technology developers can apply human factors principles to develop POC technologies efficiently, effectively, and safely for a multitude of settings. Healthcare workers can utilize human factors principles to optimize the implementation of POC technologies, allowing for increased access to diagnostic testing and ultimately for a positive impact on patient care and public health.

The Atlanta Center for Microsystems Engineered Point-of-Care Technologies (ACME POCT), part of the National Institute for Biomedical Imaging and Bioengineering’s Point-of-Care Research Network (POCTRN), has witnessed the disconnect that occurs when human factors are not considered in the development of POC technologies. As part of the National Institutes of Health’s (NIH) Rapid Acceleration of Diagnostics (RADx) initiative, ACME POCT has served as a national test validation center for meritorious COVID-19 diagnostics (Nehl et al., 2021). ACME POCT’s RADx Technology Validation Core (RADx TVC) is made up of a transdisciplinary team that includes clinical and laboratory validation testing, human factors evaluation, regulatory guidance, and engineering assessment. Of more than 50 COVID-19 POC technologies tested and assessed, several are now on the market and available for home use. The ACME POCT human factors team contributed to this evaluation process by assessing human factors in those devices that ultimately made it to the market and also in identifying other POC technologies that were at high risk of poor outcomes due to human factors concerns.

To facilitate future development of POC technologies, we leverage our experience as part of the NIH RADx initiative to review the fundamental principles of human factors in the context of POC diagnostics. For the technology developer, we describe the application of human factors principles for POC diagnostic design and the nuances of POC environments that can affect human error, all within the context of regulatory guidance. For the clinician, we describe the core human factors and usability principles in assessment of POC diagnostics. Finally, we present the “dos” and the “don’ts” of important lessons learned during this new paradigm for rapid development and evaluation of POC technologies that can be applied beyond the current pandemic.

Introduction to POC technologies
POC technologies are used at the location of patient care, which encompasses a spectrum of settings from near-patient laboratories within hospitals and emergency departments to outpatient clinics, community settings, and even patients’ homes (Külc et al., 2020). The uncontrolled environments in which these POC technologies are used and the absence of training of the intended users present significant potential for human error. As such, we frame our discussion of the application of human factors principles for reducing likelihood of errors in the context of hypothetical POC technologies.

Consider a typical POC diagnostic test kit for detecting SARS-CoV-2 illustrated in Figure 1 (not intended to represent a real product). The test kit consists of a nasal swab, lateral flow assay (LFA) device, test tube with buffer solution (called buffer tube), and dropper top. The test device has a port for inserting the sample, as well as a window where a test line and control line display test results. Figure 1 represents instructions illustrating an example protocol for this test. First, the user collects an anterior nares specimen using the nasal swab. Next, the user inserts the swab into the buffer tube and rotates the swab to mix the sample with the buffer solution. The user squeezes the sides of the buffer tube while removing the nasal swab, thereby squeezing out as much solution as possible into the buffer tube. Next, the user attaches a dropper top to the buffer tube and flicks the tube to further mix the sample with the buffer. Then the user inverts the buffer tube and squeezes five drops of solution into the test device’s sample port. After waiting 15 min, the user inspects the control line and test line to interpret results according to the instructions shown in Figure 2. If the test and control lines are both present, then the result is positive, even if the test line is very faint. If the control line is present but the test line is absent, then the result is negative. Absence of a control line renders the test invalid.

Introduction to POC settings
Each POC setting has unique information flow and environmental conditions. There may be many points of interaction among patients and healthcare providers in the information flow from sample collection to clinical decision-making. There may also be environmental elements (e.g., temperature, humidity, lighting) that vary between settings. To understand this, we suggest a simple framework that can guide the design and implementation of POC technologies. First, a POC environment might be controlled or uncontrolled. Second, a POC user might be highly trained or minimally trained.

Controlled environments with trained users often include near-patient labs in inpatient settings like intensive care units (ICUs) or emergency departments (EDs). The prototype for this scenario is the ED (Figure 3). Specimens are collected by a trained staff member, and testing is performed at or near the patient in an optimized environment. Ideally, the temperature and humidity are regulated and there is adequate space, lighting, and limited noise distraction. The user (e.g., medical technician, respiratory therapist, nurse, physician assistant, physician) has been trained and follows a standardized protocol. These settings are regulated by hospital policy and external regulatory agencies. Often, these settings are linked to a centralized clinical laboratory where a gold standard test or follow-up testing can be performed to verify POC results. The results are then interpreted by a medical professional who makes a clinical decision.

The “uncontrolled environment and minimally trained user” scenario that encompasses home use might also describe settings such as schools or workplaces where testing is performed...
by a teacher or employee without laboratory training in a facility not originally intended for medical testing to be performed. Considering household use as the prototype for this scenario, information flow for diagnostic testing must be simplified (Figure 3) and is generally linear, only encompassing one or two people (e.g., adult and child). Sample collection, preparation, testing, analysis, and interpretation of results are usually intended to be performed by one person who is guided by written instructions without additional training. There are few, if any, points of human-human interaction. Despite the simplicity of information flow in this setting, there is no regulation of the testing process or environment, and home environments can vary considerably. 

For example, there may be poor light and increased humidity. Users in a home setting are likely to experience distractions or interruptions. This variability of the home environment may influence the accuracy of testing. Additionally, self-testing while a user is ill, and therefore not at peak performance, adds another challenge to what might already be a novel experience. In the home setting, interpretation of results and medical decision-making may also be challenging. Depending upon a person’s medical literacy, the implications of a test result may not be clear. Without the guidance of a trained medical professional who can explain the meaning of a result, a person may experience uncertainty and anxiety about what to do with the information they have obtained.

During the COVID-19 pandemic, variations on “uncontrolled environment and highly trained user” scenarios have become more common. A prime example is outdoor drive-through testing sites that have been established throughout the United States, due to increased demand for testing and the safety benefits of interacting in an outdoor environment. In this setting, the flow of information is structured and there are numerous human-human interactive points. Staff are trained to collect, prepare, test, and perform data analysis. However, there is variability in who reports data and makes recommendations for a treatment plan. In outdoor settings, environmental factors such as temperature, humidity, rain, and wind can influence test results (Pollock et al., 2021). Moreover, dedicated space for testing, analysis, and reporting results is also often limited in these settings.

Environmental control and user training are two variables that must be considered in the design and implementation of POC technologies. Knowledge of human factors provides the tools for addressing these and is a needed skill for engineers and clinicians alike.

**Regulation of POC technologies**

All clinical diagnostic testing in the United States is regulated by the Centers for Medicare & Medicaid Services (CMS), Food and Drug Administration (FDA), and Centers for Disease Control and
Prevention (CDC) through the Clinical Laboratory Improvement Amendments (CLIA) (Clinical Laboratory Improvement Amendments, 2021). Under CLIA, laboratory tests are classified as “high complexity,” “moderate complexity,” and “waived” based on risk of an erroneous result, and human factors are a key component of this risk analysis (Centers for Medicare and Medicaid Services, 2019). A POC technology for home use necessitates waived status, since the environment is not controlled and the user is minimally trained. However, many POC technologies are assigned a moderate complexity designation, which requires that they must be operated in a facility with a CLIA certificate that is obtained through inspection by CMS or an accreditation agency approved by CMS (Centers for Medicare and Medicaid Services, 2018; American Society for Microbiology, 2019). As such, POC technologies encompass both waived and moderate-complexity technologies, but those with moderate complexity designation must be used in controlled environments (e.g., near-patient labs in hospitals, clinic offices, pharmacies) by highly trained users.

**INTRODUCTION TO HUMAN FACTORS IN THE CONTEXT OF POC DIAGNOSTICS**

**Human factors basics**

The field of human factors is concerned with understanding human capabilities and limitations and applying that knowledge to the design of systems and technologies (HFES, 2021). For effective human factors design, one must first understand the user’s task and the steps required to complete that task. Second, one must understand the user’s abilities, including physical, sensory, and cognitive. Finally, one must understand the environment and how that impacts the user and the task demands. Figure 4 shows these components as a system, with the user’s ability and the task demands feeding into the environment and context, and failure or success of the task as the outcome. If the user has the ability to complete the task demands within the context, the design is usable for individuals with similar abilities and limitations. If the user’s ability does not meet the task demands within the context, the design is not usable. Figure 4 illustrates examples of human factors considerations that contribute to successful and unsuccessful usability in design. Human factors consideration during design ensures that the technology meets the user’s needs at the user’s ability level in the context where the device is being used.

To further illustrate this concept, consider the typical POC rapid antigen test that features a window that displays results (Figure 1). One line may indicate a negative result or an invalid result, and two lines indicate a positive result. The task of interpreting the test result requires the user to identify the presence of one line versus two lines, and these lines may present visually with varying intensity. Now consider a user who has reduced visual acuity or reduced contrast sensitivity. If they are performing the POC test outdoors in sunlight, they will likely be able to correctly discern the number of lines and therefore the correct result. Since their visual acuity in the bright sunlight can meet the task demands, they can successfully determine the result. However, consider the case where the test is conducted outdoors when the sky is overcast, it is raining, or it is nighttime. The user’s reduced vision, coupled with the dim lighting, may prevent the user from discerning the number of resulting lines. Because the task demand exceeds the user’s visual acuity in this context, the result is not discernible, and by extension, the device is not usable.

**Human errors or user errors**

There are several types of human errors, also known as user errors. One subgroup of errors is mistakes, which are also referred to as thinking errors (Reason, 1990). These are more likely to occur with novice users and are categorized as rule-based or knowledge-based. A rule-based error might occur when a user applies know-how from another situation to a novel situation. For example, a home user has received nasopharyngeal swabs from healthcare providers in the past and incorrectly assumes that is the appropriate depth to swab, rather than the instructed 3/4 inch (Figure 5). A knowledge-based error might occur when a home user opens the swab packaging from the swab tip side rather than the handle side because they are not aware that contamination may occur if they touch the swab tip (Figure 5).

Another subgroup of errors is referred to as skill-based errors. These tend to occur with users who are trained in using the device in question; that is, they have all the knowledge and training needed to successfully conduct the testing. Skill-based errors include slips of action and memory lapses (Reason, 1990). An
example of a memory lapse is a tired healthcare worker who for-
gets to flick the buffer tube to mix the sample and buffer before
adding the solution to the test device, though he typically executes
the protocol with no lapses (Figure 5). An example of a slip of action
is an experienced healthcare worker who loses count when
squeezing the drops of solution onto the test device (Figure 5).

Design-induced errors or latent human errors
Design-induced errors, also known as latent human errors, are er-
rors that can be attributed to poor design of the product itself
(Reason, 1990). Design-induced error should be largely avoidable
with adherence to sound design principles. In the case of our
generic test device, a single control line indicates a negative result,
while a single test line without a control line indicates an invalid
result. If the lines are not labeled on the device, then users
observing a single line may confuse negative results with invalid re-
results. Clearly labeling the location of each of these lines on the test
device will aid the user in correctly interpreting the result (Figure 5).

However, even by integrating human factors design principles,
it is not possible to eliminate all errors. The goal is to reduce the
frequency of errors, eliminate sources of design-induced error, and minimize the level of harm when errors do occur.

How do we reduce the frequency of er-
rors in POC testing? Human factors
analytical methods can be utilized to
minimize sources of error during the plan-
ning and design phases of technology development as well as to
validate technologies after development. Formative evaluations
(also referred to as preliminary analyses or evaluations) are
testing methodologies applied during the development process,
and are intended to identify potential sources of error in the
design. Formative evaluations include task and error analyses,
design failure mode and effects analyses (DFMEAs), heuristic
evaluations, expert reviews, cognitive walkthroughs, prototype
testing, and simulated use studies. Summative evaluations are
testing methods intended to be used at the end of a design cycle
and are also referred to as validation testing. These evaluations
are intended to determine whether a device can successfully
be used by all intended user groups, under all expected condi-
tions, without serious errors or harm to the user occurring. Sum-
mative evaluations typically consist of extensive user testing un-
der either actual or simulated conditions. If a device does not
pass the predefined criteria, it should go through a second iter-
anation of the design process and the summative evaluation should
be repeated.

In 2016, the U.S. Food and Drug Administration (FDA) issued a
guidance document entitled “Applying Human Factors and Us-
ability Engineering to Medical Devices” (FDA, 2016).
The guidance document provides an excellent overview of the role of human factors and usability engineering in risk management. It also reviews a wide array of techniques for assessing the usability of a medical device. In the following sections, we introduce a sampling of assessment techniques for evaluating the usability of POC technologies: DFMEA and heuristic analysis.

### Design failure mode and effects analysis

DFMEA is an analytical tool for quantifying the risk level of a design. Let us walk through the process of conducting a DFMEA using the example of a nasal sample collection kit (Table 1). First, the kit is divided into components. For example, the sample collection kit consists of a swab handle, swab tip, packaging, and instructions for use. Next, for each component, all possible functions are defined. For example, the functions of the packaging are to protect the swab tip from contamination and to give the user access to the swab. Then for each function, failure modes are identified; these are distinct ways in which a component could fail to meet its function. For instance, the packaging may not be openable by the user due to the force required to tear the packaging exceeding the functional abilities of the user, or the packaging may expose the swab tip (rather than handle) when opened.

The next step in the DFMEA is to identify possible effects of each failure mode. If the packaging is not openable by the user, then the sample collection procedure cannot be conducted. If the open package gives access to the swab tip but not the handle, then the user may touch the swab tip and contaminate it. The user may recognize this risk and throw away the swab, or they may collect the sample using a contaminated swab that could interfere with the assay or introduce a pathogen from the environment. For each effect, the severity may be rated from 1 (no effect) to 10 (catastrophic) (Carlson, 2012). The unopenable packaging has a severity of 8, corresponding to a loss of the testing kit’s primary function. The severity of using a contaminated swab on a patient is a 10, corresponding to a failure to meet safety or regulatory requirements.

In the next step of the DFMEA, the severity rating is multiplied by the occurrence rating to arrive at a quantified risk level. In the above case of unopenable packaging, the risk level is 64. For components exceeding a particular risk threshold, design changes or risk mitigation plans are recommended to reduce the severity or frequency of failure. For instance, the packaging could be redesigned with a tab that makes it easier to open, even for users with reduced dexterity. With the redesign for easy-open packaging, the occurrence of unopenable packaging would drop to 4. The risk level is then re-assessed for an updated design. Although time intensive, a DFMEA provides a comprehensive view of how designs are most likely to fail.

### Heuristic analysis

In contrast with the resource-intensive DFMEA, heuristic analysis is a time- and cost-effective method for identifying impediments to usability. In heuristic analysis, the overall design is evaluated against several usability heuristics, or rules of thumb for usable design. Jakob Nielsen’s 10 usability heuristics as outlined below are frequently used in human factors and usability evaluations (Nielsen, 1994b):

1. Visibility of system status: users should be able to discern the status of the system via feedback and display. It should be clear what available actions are.
2. Match between system and real world: the system should follow real-world conventions and use the language of the user. User actions should lead to expected outcomes.

3. User control and freedom: users should have control of the system and be able to easily “undo” actions and navigation.

4. Consistency and standards: consistent terminology, style, and actions should be used throughout the system.

5. Error prevention: the system should be designed in a way that most errors are completely avoided.

6. Recognition rather than recall: the system should minimize cognitive load on the user where possible, via use of visible instructions, menus, and examples.

7. Flexibility and efficiency of use: the system should allow for users to customize shortcuts and use experiences to accelerate their performance relative to their experience levels.

8. Aesthetic and minimalist design: the system should present only relevant information required to complete tasks in order to avoid overloading the user.

9. Help users with errors: error messages should be clear and informative, and provide enough information for users to recover from the error.

10. Help and documentation: instructional materials should be provided and should be easily searchable and focused on executable tasks.

To conduct a heuristic analysis, evaluators inspect the device or system for usability problems. For each issue, they note which component of the system is involved, which heuristics are violated, and the severity of the issue. Consider again, the example of a nasal sample collection kit that has difficult-to-open packaging. The component involved is the physical packaging. This problem violates the “error prevention” heuristic, because the packaging was not designed in a way that avoids error. In other words, this is a design-induced error. Next, a rating is determined for the severity of the issue. Nielsen lists the following severity categories (Nielsen, 1994a):

- Low: superficial or cosmetic issue; resolving would have little impact on usability.
Table 1. Example of a design failure mode and effects analysis (DFMEA) for the swabbing components of a diagnostic test kit

| Component          | Function                          | Potential failure mode                                      | Potential effect(s) of failure                          | Severity | Potential cause(s) of failure | Occurrence | Risk level | Recommendations                                                                 |
|--------------------|-----------------------------------|-------------------------------------------------------------|--------------------------------------------------------|----------|------------------------------|------------|------------|--------------------------------------------------------------------------------|
| Swab handle        | Protect sample collector from contamination | Handle is too short, causing the sample collector to touch the patient's face | The sample collector’s hand becomes contaminated | 10       | Contamination                | 3          | 30         | Depending on the depth of the required swab, make sure the handle is long enough to not crowd the face while swabbing |
|                    | Provide swab control to the user  | Material snaps while in use                                  | Sample collection is unable to proceed                 | 8        | Material Failure             | 2          | 16         | Use flexible and sturdy material                                                 |
| Swab tip           | Transport sample to buffer/device | Swab head is too large for insertion into buffer/device      | Sample is not tested properly                          | 7        | Component Incompatibility    | 2          | 14         | Find compatible swabs or use larger buffer tubes                                |
|                    | Collect sample from patient       | Swab material loses integrity and comes apart during use    | Improper sample collection                             | 7        | Material Failure             | 2          | 14         | Use synthetic swabs in place of cotton swabs                                     |
| Swab packaging     | Provide access to the swab        | Package is unopenable by user                                | Sample collection is unable to proceed                 | 8        | Accessibility                | 8          | 64         | Provide tabs and use easy-to-open material                                     |
|                    | Provide access to the swab        | Open package gives access to the swab tip instead of the handle | Swab tip is contaminated, potentially infecting patient | 10       | Accessibility or User Error  | 3          | 30         | Design opening point to be at the handle of the swab                           |
Perspective

- Medium: minor usability issue; resolving is lower priority but would improve usability.
- High: major usability issue; resolving is high priority and will greatly impact usability.
- Critical: usability catastrophe; resolving is highest priority, and product should not be used until resolved.

In the case of difficult-to-open packaging, the severity rating would be high, as this issue impedes users from being able to conduct the test or prompts the user to introduce an unsafe tool such as a sharp knife in order to open the packaging.

When conducting a heuristic analysis, a single evaluator will identify a fraction of the usability problems (Nielsen and Molich, 1990). By including 3–5 independent evaluators, 60%–75% of usability problems are identified (Zhang et al., 2003). All the usability problem lists are then combined, and each evaluator independently rates the severity of each usability problem on the comprehensive list. By averaging these ratings, the mean severity rating is determined.

INTEGRATION OF HUMAN FACTORS IN THE DESIGN OF POC TECHNOLOGIES

We often hear and repeat the maxim that diagnostic tests should be “as simple as a digital pregnancy test,” but in reality, what constitutes design simplicity? Is it the method of sample collection, low number of components, minimalist user interface, clear communication of errors? Is it the time from sample collection to result or perhaps ease of result interpretation on the digital readout? There are numerous seemingly complex elements incorporated into the design of a digital pregnancy test; however, the integration of the individual elements and consideration of human factors results in a simple-to-use test for most people. It should be noted that not all digital pregnancy tests are necessarily usable. Like all diagnostic tests, error-prone designs exist, but when done well, these tests have the potential to be simple, effective, and highly usable.

When considering the incorporation of human factors into the design of POC technologies, the initial step is to identify the intended users and identify their capabilities and limitations (Figure 6). Users of POC diagnostic tests are likely different from users of lab diagnostic tests. For laboratory tests, we can assume that the lab technicians have knowledge of lab-related terminology and techniques. However, the knowledge and skills of POC test users as a group are likely to be far more varied. POC test users could range from airport staff to sporting event ushers to pharmacists to nurses in clinics. Their functional abilities may range from highly capable to disabled. Furthermore, there may be multiple users involved in the performance of a single POC diagnostic test; for instance, a patient may collect a sample themselves, while a different person processes the test and interprets the results. Designers need to ensure that the POC device is designed to enable the full range of users to perform each step of the diagnostic test successfully.

Next, the environment in which the POC technology will be used should be considered in the design (Figure 6). Tests can be conducted in outdoor parking lots, airport terminals, sports stadiums, concert halls, and even patients’ homes. Outdoor testing environments present a particular challenge due to the variability in weather conditions. It is not sufficient to design a test for a single testing environment or a single set of testing conditions. Today’s POC diagnostic tests must meet the demands of testing across a wide range of humidity, temperature, noise, and lighting conditions.

The next step is to define the use cases or scenarios (Figure 6). A use case includes all the tasks associated with the use of the device as well as the environmental variables that may impact performance. A use case should also consider all users involved in any part of the scenario. Use cases should cover typical use scenarios as well as any atypical scenarios that might be anticipated. In the case of our POC diagnostic test, a use case could be defined for a single, experienced user administering tests in a clinic as well as two novice users (perhaps one is collecting the sample and one is conducting the tests) in an outdoor, temporary testing site.

The last step is to define the user interface (Figure 6). The user interface consists of any part of the system that the user will interact with. This includes the device itself (the hardware), the packaging the device comes in, any instructional or training materials, and any associated software. Defining the user interface is an important part of the design process, as error mitigations applied to portions of this interface are typically the most effective method to avoid user errors. For our POC diagnostic test, the user interface may consist of the entire test kit (swab, buffer tube, test device, and all packaging materials), the instructional materials (an instructions-for-use document and a quick reference guide), and a training video created by the device company.

The American National Standards Institute (ANSI) and the Associate for the Advancement of Medical Instruments (AAMI) released a guidebook of human factors principles for the design of medical devices (Clinical Laboratory Improvement Amendments, 2009). This document details general principles as well as principles for specific aspects of user interfaces and is a useful reference for incorporating human factors into the design of medical devices.

Incorporating usability heuristics into design

In the “heuristic analysis” section, ten usability heuristics were introduced that are used in conducting heuristic evaluation. Not only do these heuristics aid evaluators in the assessment of POC technologies, but they may also serve as usability rules of thumb to incorporate into design.

Visibility of system status

Systems must be designed to convey to users the status of the system via feedback and display. The system must also make it clear to users what actions are available. Consider the prototype LFA test (Figure 1). When the user attaches the dropper top to the buffer tube, a physical “click” into place indicates to the user that they have successfully attached the top and the task is complete.

Now, consider designing a POC test consisting of an instrument with a touchscreen that processes single-use test cartridges. First, the design of the physical instrument should have a slot that is the right size to fit the cartridge. It must be
designed it in such a way that when the user sees or feels this slot, they can tell that the proper physical interaction is to insert a test cartridge with a specific orientation. Now consider the instrument’s touchscreen display. When the user inserts the cartridge, the instrument’s user interface (UI) should display feedback that the user has inserted the cartridge. Later, once the user initiates processing, the UI could display the text “Processing” or provide a countdown timer in order to provide clear feedback to the user about the system status (Figure 7). Alternatively, a simple status indicator light that informs the user of the current system status may provide adequate feedback.

**Match between system and real world**

The system must be designed to follow real-world conventions and to use the language of the user at the user’s literacy level. It should ensure alignment between the system and user actions. For example, language that is familiar to the intended user should be used in instructional materials. For the home use case, materials should use plain language at the eighth-grade literacy level and avoid medical or scientific jargon (FDA, 2001). Match real world conventions where possible, such as use of color as a redundant coding mechanism to emphasize messages in the materials (Figure 7).

**User control and freedom**

A well-designed system must be easy to navigate, give users control, and let users “undo” actions. To examine this heuristic in practice, consider a PCR system to detect the SARS-CoV-2 virus. Processing is often triggered by interaction with a UI screen. To prevent error, the UI should require the user to confirm critical actions, such as a pop-up screen that asks users to “confirm sample has been inserted” before starting processing. Additionally, “back” buttons assist users in navigating through various UI screens (Figure 7).

**Consistency and standards**

There should be consistent terminology and style throughout a system, including the instructions and documentation. For example, when writing instructions, do not use synonyms to refer to a test component. Users will get confused if a single component is sometimes referred to as a “vial” and sometimes as a “buffer tube” and may wonder if these are two separate components. Components should be referred to consistently throughout materials (Figure 7).

**Error prevention**

In the “human errors or user errors” and “design-induced errors or latent human errors” sections, we introduced user error (human error) and design-induced error (latent human error). It is not possible to design a system to eliminate all errors, but the system can be designed in such a way that most errors are completely avoided. For example, when designing a test device, the number of ports for inserting solution should be considered. If there are multiple ports, users may insert a specimen into the wrong port. By designing the device with only a single port, this error can be prevented (Figure 7).

**Recognition rather than recall**

The system should minimize cognitive load (e.g., translating units of measurement, holding multiple steps in working memory) on the user where possible, via use of visible instructions, menus, and examples. Consider an LFA test with two result lines; the presence or absence of the test line indicates the result, and the presence of the control line indicates that the test is valid.
Figure 7. Examples of Nielsen’s 10 Usability Heuristics

The figure above illustrates how Nielsen’s Usability Heuristics can be incorporated in a POC diagnostic test. Panels 1, 3, 7, 8, and 9 all illustrate examples of user interface screens that adhere to the corresponding heuristic. Panels 2, 4, and 10 all illustrate how instructional materials should also be considered when conforming to Nielsen’s Usability Heuristics in design. Lastly, panels 5 and 6 show LFA device comparisons where the device at the top of each illustration follows the corresponding heuristic.
If the line positions are not labeled on the device, it takes mental work to determine which line appears and recall whether that line is the test line or the control line. Labeling the two lines reduces users’ cognitive load (Figure 7).

**Flexibility and efficiency of use**

The system should allow users to customize shortcuts and experiences to accelerate their performance relative to their experience levels. Consider an Internet-of-Things (IoT) POC device that connects to a mobile application. For first-time users or one-time users, it may be helpful to view a training video. However, an experienced user would likely get frustrated if the training video played every time the app is opened. By providing a “skip” function to allow experienced users to bypass viewing the training video, apps can give users the flexibility to customize their experience (Figure 7).

**Aesthetic and minimalist design**

The system should present only relevant information required to complete tasks in order to avoid overloading the user. For example, when designing a real-time PCR test for detecting SARS-CoV-2 in POC settings, the display results may be programmed to read “Positive” or “Negative” without displaying the raw data onscreen (Figure 7).

**Help users with errors**

Error messages should be clear and informative and provide enough information for users to recover from the error. For example, an LFA device should be designed so that an invalid result is visually interpretable by the user, and follow-up steps should be provided in the instructions for use (IFUs). This communicates to the user that an error has occurred (whether that error is design-induced or due to human error) and provides actionable steps to resolve the error.

Next, consider a POC test comprising a cartridge that is inserted into an instrument. Suppose the user inserts the cartridge upside down, and the instrument displays a warning: “Error: cannot process test.” This error message is not useful because it does not inform the user what went wrong or how to fix it. Instead, include a more explicit error message, such as: “The cartridge has been inserted incorrectly. Please reinsert cartridge with arrows pointing down” (Figure 7).

**Help and documentation**

Instructional materials should be designed to be easily searchable and focused on executable tasks. Instructional materials should cover every task that a user is expected to execute and use language familiar to the user. For example, a home-use test should have instructions written at an eighth-grade reading level without medical jargon and should include a description of all tasks, even those that are second nature to medical professionals (FDA, 2001) (Figure 7).

**LESSONS LEARNED FROM THE NIH RADx INITIATIVE—THE DOS AND DON'TS OF POC TECHNOLOGY DESIGN**

Through the NIH’s RADx initiative, the ACME POCT evaluated more than 50 different COVID-19 diagnostic test systems. During the evaluation process of these diagnostics, consideration of human factors was repeatedly evaluated and discussed. Some technologies excelled in the incorporation of human factors principles, whereas others poorly incorporated human factors into their design. To assist in the development of future POC technologies, we make the following recommendations for technology developers and clinicians to consider during design and development of POC technologies. Table 2 includes an abbreviated list of some of the most commonly encountered “dos and don’ts” of POC technology design.

**Sample collection**

**Make it easy for the user to properly collect the sample required for the POC technology**

Two main types of samples are used for COVID-19 POC tests: nasal swab and saliva. For both sample types, the sample collection method and equipment significantly impact the system’s usability. Two potential sources of error with saliva collection are if the user cannot physically collect the saliva or if the user cannot determine the necessary volume of saliva needed. Rather than having a patient spit directly into a small vial, include a funnel or sponge-type swab for saliva collection. If expectorated saliva is to be collected, clear indicators of the required volume should be provided. For example, include a high-contrast fill line on a vial to mark the minimum volume level. Saliva-based tests are susceptible to a further pitfall: the test may be invalidated if the patient eats, drinks, smokes, chews gum, or brushes their teeth prior to sample collection. It is therefore essential to include a clear warning in instructional materials regarding pre-test restrictions. When designing an ideal system, we would include provisions for notifying a patient in advance to refrain from these activities. At a minimum, the instructional materials should include a warning immediately preceding the procedure for collecting the saliva sample.

Instructional materials are also paramount to the usability of nasal sample collection. It is essential to provide clear instructions on how to properly collect the sample, such as how deep to insert the swab, how to know when you have inserted it to the right depth, and, once inserted, how many rotations to circle the swab within the nostril. Many at-home users will not know how far back to take a nasal swab, especially with many news articles indicating that samples are taken from the back of the nose. Diagrams are key. Furthermore, if the person collecting the sample is not a medical professional, translate jargon like “mid-turbinate” and “anterior nasal” into terms that are easier to understand (e.g., “insert the entire soft tip of the swab into one nostril, no more than \(\frac{3}{4}\) inch, until you feel slight resistance”). Finally, keep in mind that the person who collects the sample may not be the person who conducts the rest of the diagnostic test.

**Sample preparation**

**The more sample preparation steps, the more opportunities for error—Minimize the steps and streamline the workflow where possible**

When preparing the sample, one frequent source of error is in measuring solutions. To mitigate measurement error, it is ideal to package pre-measured volumes of solutions; it is easier for...
Table 2. Dos and don’ts for POC test design/what to look for when evaluating a test

| Dos                                      | Don’ts                                                                 |
|------------------------------------------|------------------------------------------------------------------------|
| Instructional materials                  |                                                                         |
| ✓ Multiple formats                       | ✗ Small font                                                           |
| ✓ High-quality images                    | ✗ Few or no images                                                     |
| ✓ Minimum 14 pt font                     | ✗ Unnumbered steps                                                    |
| ✓ Numbered steps                          | ✗ Inconsistent terminology                                             |
| ✓ Audience-appropriate language          | ✗ Warnings listed only in beginning of instructions                   |
| ✓ Cohesive, comprehensive               |                                                                        |
| ✓ Troubleshooting guide                   |                                                                        |
| Physical considerations                  |                                                                         |
| ✓ Easy-to-open packaging with notches or overhanging tabs | ✗ Packaging difficult to open                                         |
| ✓ Components large enough to grasp easily | ✗ Small fonts                                                          |
| ✓ Buffer tube racks/tubes that stand on their own | ✗ Low contrast                                                        |
| ✓ Easy-to-open caps with grips           | ✗ Small components                                                    |
| ✓ Locking mechanism                      | ✗ Color-only indicators                                                |
| Protocol                                 |                                                                         |
| ✓ Pre-measured volumes of solutions      | ✗ Requiring users to measure solutions                                |
| ✓ Pre-mixed solutions when possible      | ✗ Requiring users to mix solutions                                    |
| ✓ Fewer components                       | ✗ Transferring samples or solutions from one component to another      |
| ✓ Minimal number of steps                |                                                                        |
| Process indicators                        |                                                                         |
| ✓ Initiating processing is clear         | ✗ Leaving users uncertain through lack of feedback                     |
| ✓ Progress indicators and time expecta-  |                                                                        |
|   tions                                  |                                                                        |
| ✓ Indication when processing is complete |                                                                        |
| Result interpretation                     |                                                                         |
| ✓ Unambiguous digital “positive,” “neg- | ✗ Requiring users to make judgment calls for result interpretation      |
|   ativé,” and “invalid” text read-        | ✗ Color-only interpretation                                            |
|   outs                                   | ✗ Results lines too close to each other                                |
| ✓ Labeled control and test lines for    |                                                                        |
|   non-digital tests                      |                                                                        |
| ✓ Images of possible results in           |                                                                        |
|   instructional materials                |                                                                        |

a user to add the contents of a whole vial rather than measure out a specific volume. If some measurement is required, a fixed volume pipette is preferable. Error can also occur when handling multiple test tubes or mixing multiple solutions together. Provide test tube holders or stands to prevent spilling, and when possible, provide pre-mixed solutions.

Sample processing

From start to finish, visibility of system status is central to a user’s experience—Once processing has started, feedback is key

First, the system should make it clear how to initiate processing. For instance, an electronic test could have a “Start” button. Progress indicators are a good way to inform the user that something is happening, and users also benefit when systems provide processing time expectations. Finally, the system should provide clear indication when processing is complete.

Result interpretation

Explicit readouts reduce the risk of error by taking the guesswork out of result interpretation

Most COVID-19 tests indicate results in one of three ways: (1) color change, (2) control and test lines, or (3) digital “positive,” “negative,” or “invalid” readouts. Color interpretation is rife with potential for errors, and it is particularly difficult in environments with dim lighting and for users with reduced vision. For LFAs featuring two lines, common pitfalls include insufficient delineation between the control line and test line and difficulty discerning the presence of the lines. To maximize ease of line interpretation, label the test and control lines directly on the physical device. Also, include in the documentation images of all possible results; this is especially important for clarifying results in the event of a faint test line. Note that faint test lines could lead to errors, and faint lines are particularly challenging to discern in dimly lit environments and for users with reduced vision. With regard to human factors, the ideal solution is to provide unambiguous results in the form of digital “positive,” “negative,” and “invalid” readouts.
Instructional materials
Create multiple types of instructional materials—Use pictures, large fonts, and clear wording to make instructions easy to read

Detailed IFUs provide extensive introductions to new users, and they are useful reference materials for experienced users troubleshooting or performing infrequent maintenance procedures. Using a readability calculator, assess the instructional materials to aim for an eighth-grade reading level (FDA, 2001). For experienced or subsequent users, a 1- or 2-page quick reference guide (QRG) is useful. Users can walk through the QRG every time they conduct the testing procedure, without being hindered or frustrated by the details of a longer IFU. Alternate formats, such as an instructional video, may be appropriate for some POC technologies and increase the accessibility of the system. Instructional materials should use a minimum font size of 14 pt to ensure readability. To maximize usability, number the steps and include a picture for every step of the procedure in both the IFU and QRG. Also include labeled pictures of all POC test system materials in the IFU. Pictures should be realistic or actual images, and they must have strong contrast. Unfortunately, many IFUs and QRGs utilize small fonts, pictures with too little contrast, and too few pictures. IFUs should be comprehensive and make it clear what steps a user must perform to complete the procedure successfully. Sometimes instructions use jargon or fail to give enough detail, requiring the user to rely on outside knowledge to figure out how to conduct the procedure.

CONCLUSIONS

The COVID-19 pandemic has taught the medical community, diagnostics industry, and the public that POC testing is here to stay and, if anything, will play an ever-increasing role in disease diagnosis away from centralized clinical laboratories. While this will certainly bring clinical value, increased quality of life, and potential improvement in patient outcomes, POC testing in these settings will also expose challenges and limitations that need to be overcome. In particular, the importance of integrating human factors considerations into the development of POC diagnostic technologies from the very beginning of the ideation and design process cannot be overstated. This will not only enable biomedical engineers and diagnostics developers to efficiently define the clinical use case, regulatory strategy, and manufacturing plans for their proposed technology but also streamline the pathway to intended use—ultimately allowing enhanced delivery of medical care and public health interventions in a wide spectrum of settings.

ACKNOWLEDGMENTS

This work is supported by NIH grants U54 EB027690 02S1, U54 EB027690 03S1, U54 EB027690 03S2, UL1 TR002378. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

American Society for Microbiology (2018). Bringing the Laboratory Test to the Patient with Point-of-Contact Tests. https://asm.org/Articles/2018/October/Bringing-the-Laboratory-Test-to-the-Patient-with-P.
American National Standards Institute (2009). Human Factors Engineering-Design of Medical Devices in ANSI/AAMI HE75:2009 (Association for the Advancement of Medical Instruments).
FDA (2016). Applying Human Factors and Usability Engineering to Medical Devices: Guidance for Industry and Food and Drug Administration Staff, CDRH (U.S. Food and Drug Administration).
Carlson, C. (2012). Effective FMEAs: Achieving Safe, Reliable, and Economical Products and Processes Using Failure Mode and Effects Analysis, 1 (John Wiley & Sons).
Centers for Medicare & Medicaid Services (2019). How to Obtain a CLIA Certificate (U.S. Department of Health & Human Services).
Chowell, G., Dahal, S., Bono, R., and Mizumoto, K. (2021). Harnessing testing strategies and public health measures to avert COVID-19 outbreaks during ocean cruises. Sci. Rep. 11, 15482. https://doi.org/10.1038/s41598-021-95032-4.
Clinical Laboratory Improvement Amendments (2021). https://www.fda.gov/medical-devices/vd-regulatory-assistance/clinical-laboratory-improvement-amendments-clia.
FDA (2001). Guidance on Medical Device Patient Labeling; Final Guidance for Industry and FDA Reviewers (Center for Devices and Radiological Health).
Gubbins, P.O., Klepser, M.E., Dering-Anderson, A.M., Bauer, K.A., Darin, K.M., Klepser, S., Matthias, K.R., and Scarsi, K. (2014). Point-of-care testing for infectious diseases: opportunities, barriers, and considerations in community pharmacy. J. Am. Pharm. Assoc. 54, 163–171. https://doi.org/10.1331/japa.2014.13167.
HFES (2021). What Is Human Factors and Ergonomics? (About HFES).
Kennedy, C., Yeh, P.T., Johnson, C., and Baggaley, R. (2017). Should trained lay providers perform HIV testing? A systematic review to inform World Health Organization guidelines. AIDS Care 29, 1473–1479. https://doi.org/10.1080/09540121.2017.1317710.
Klic, T., Weissleder, R., and Lee, H. (2020). Molecular and immunological diagnostic tests of COVID-19: current status and challenges. iScience 23, 101406. https://doi.org/10.1016/j.isci.2020.101406.
Kim, S.J., Watson, K., Khare, N., Shastri, S., Goia Pintpo, C.D., and Nazir, N. (2021). Addressing racial/ethnic equity in access to COVID-19 testing through drive-thru and walk-in testing sites in Chicago. Med. Res. Arch. 9, 2430. https://doi.org/10.18103/mra.v9i6.2430.
Murray, E.T., Fitzmaurice, D.A., and McCahon, D. (2004). Point of care testing for INR monitoring: where are we now? Br. J. Haematol. 127, 373–378. https://doi.org/10.1111/j.1365-2414.2004.05154.x.
Nehi, E.J., Heilman, S., Ku, D., Gottfried, D., Farmer, S., Mannino, R., Tyburiski, E., Sullivan, J., Suesssmith, A., Passit, L., et al. (2021). The RADx tech test verification core and the ACME POCT in the evaluation of COVID-19 testing devices: a model for progress and change. IEEE Open J. Eng. Med. Biol. 2, 142–151. https://doi.org/10.1109/omjeb.2021.3070825.
Nielsen, J. (1994a). Severity Ratings for Usability Problems. https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/.
Nielsen, J. (1994b). Enhancing the explanatory power of usability heuristics. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp. 152–158. https://doi.org/10.1145/191666.191729.
Nielsen, J., and Molich, R. (1990). Heuristic evaluation of user interfaces. Proc. ACM Chi’90 Conf., 249–256. https://doi.org/10.1145/97243.97281.
Parcell, B.J., Brechin, K., Allstaff, S., Park, M., Third, W., Bean, S., Hind, C., Farmer, R., Chandler, D., and Chalmers, J.D. (2020). Drive-through testing for SARS-CoV-2 in symptomatic health and social care workers and household members: an observational cohort study. Thorax 75, 1109–1111. https://doi.org/10.1136/thoraxjnl-2020-215128.
Pollock, N.R., Jacobs, J.R., Tran, K., Cranston, A.E., Smith, S., O’Kane, C.Y., Roady, T.J., Moran, A., Scarry, A., Caroll, M., et al. (2021). Performance and implementation evaluation of the abbott BinaxNOW rapid antigen test in a high-throughput drive-through community testing site in Massachusetts. J. Clin. Microbiol. 59, e00083-21. https://doi.org/10.1128/jcm.00083-21.

Price, C.P. (2001). Regular review: point of care testing. BMJ 322, 1285–1288. https://doi.org/10.1136/bmj.322.7297.1285.

Rayment, M., Thornton, A., Mandalia, S., Elam, G., Atkins, M., Jones, R., Nar done, A., Roberts, P., Tenant-Flowers, M., and Anderson, J.; Ann K. Sullivan on behalf of the HINTS Study Group (2012). HIV testing in non-traditional settings—the HINTS study: a multi-centre observational study of feasibility and acceptability. PLoS one 7, e39530. https://doi.org/10.1371/journal.pone.0039530.

Reason, J. (1990). Human Error (Cambridge University Press).

Robinson, J.H. (2016). Bringing the pregnancy test home from the hospital. Soc. Stud. Sci. 46, 649–674. https://doi.org/10.1177/0306312716664599.

Tromberg, B.J., Schwetz, T.A., Perez-Stable, E.J., Hodes, R.J., Woychik, R.P., Bright, R.A., Fleurence, R.L., and Collins, F.S. (2020). Rapid scaling up of covid-19 diagnostic testing in the United States — the NIH RADx initiative. N. Engl. J. Med. 383, 1071–1077. https://doi.org/10.1056/nejmsr2022263.

Zhang, J., Johnson, T.R., Patel, V.L., Paige, D.L., and Kubose, T. (2003). Using usability heuristics to evaluate patient safety of medical devices. J. Biomed. Inform. 36, 23–30. https://doi.org/10.1016/s1532-0464(03)00060-1.