



**NCSBN**  
Leading Regulatory Excellence

## ***Past Event: 2021 NCSBN REx-PN Conference - Sensitivity & DIF Video Transcript***

©2021 National Council of State Boards of Nursing, Inc.

### **Event**

Past Event: 2021 NCSBN REx-PN Conference

More info: <https://www.ncsbn.org/16034.htm>

### **Presenters**

Latrice Johnson, MSN, RN, Manager, Content Processes, Examinations, NCSBN

Hong Qian, PhD, Manager, Measurement and Testing, Examinations, NCSBN

- [Latrice] Hello. My name is Latrice Johnson, RN manager of content development here at NCSBN. During my 12 years of test development, I have been actively involved in numerous item development panels. Today, Dr. Hong and I will be talking about sensitivity review and differential item functioning, otherwise known as DIF, two procedures that include item review panels that we host to reduce bias. By the end of this presentation, you will be able to identify the purpose of conducting the sensitivity review panel and the DIF review panel. You will understand why these panels are necessary and what it means for an item to DIF.

Finally, you will be able to describe how NCSBN conducts the sensitivity and DIF review panels. Before we discuss sensitivity and DIF, let's look at the REx-PN item development process. Review of this process will help you identify the many actions taken to develop the REx-PN items. This diagram illustrates the steps used in the development of the REx-PN exam.

The REx-PN item development process is a multifaceted process adhering to the standards outlined in professional test development. The exam process starts at the top with nursing practice. Nursing practice provides the foundation for the REx-PN item development process by identifying activities that can be performed by an entry-level nurse.

I'll briefly review each step within the item development process. To summarize, after the practice analysis and test plan development, items are developed by volunteer nurse educators at item writing panels. Then, items are revised by our content developers here at NCSBN to ensure that the items look as if one writer has written them since we do have many volunteers who attend our panels.

Items are then submitted for additional editorial review, which ensures that items are grammatically correct. After editorial review, items are sent to the item review panel. Volunteer nurse clinicians who

work directly with entry-level nurses attend item review panels to ensure that the items are accurate, entry-level, current practice, and linked appropriately to the test plan.

Items are then sent to sensitivity review. I'll share more details about sensitivity review later. After sensitivity review, items then go on to the exam for pre-testing. We use pre-testing for data collection. These items do not count towards the candidate score. Pre-test items and items that are operational or count towards the candidate score may undergo differential item functioning or DIF review, which Dr.

Hong will discuss later. Items are also reviewed by nursing regulatory bodies to determine if the content is appropriate for their province. And the last step is exam administration, of which items are now operational and count towards the candidate score. As you can see here, the entire process is rigorous and requires a lot of nurse volunteers.

It takes approximately 18 months to 2 years for an item to be developed and become operational on the exam. Now, let's go back to sensitivity and DIF review. What is the purpose of sensitivity and DIF reviews? Do you know that the REx-PN will be administered to candidates who plan to practice in British Columbia and Ontario?

These candidates represent global diversity. We want to make sure that the exam is fair to all candidates. This is the purpose of sensitivity and DIF reviews. It is critical that every candidate has a fair opportunity to demonstrate their nursing knowledge, skills, and abilities. Fairness in an exam means that the exam provides an equal opportunity for all candidates.

We want our examination to make distinctions between people based on their knowledge, skills, and abilities within the nursing profession alone and nothing else. So, how do we ensure fairness on REx-PN? As I mentioned before, there are two procedures during the item development process to ensure fairness.

The first one is sensitivity review panel. This panel includes a diverse group of volunteer laypersons and two nurses. The panelists review newly written items before the items are used for pre-testing. The second procedure is DIF review, which Dr. Hong will discuss next. For now, I want you to notice that each panel includes volunteer laypersons from diverse backgrounds.

Let's talk about the sensitivity review panel. During the sensitivity review panel, the participants are asked to review each item to identify inappropriate language, underlying assumptions and stereotyping, ethnocentrism, and inflammatory language or content. Now, let's take a closer look at our sensitivity review criteria.

Inappropriate language. Our panelists are instructed to flag any item that classifies a client by their diagnosis, age, religion, or disability. They are also instructed to flag items that use expressions and slang that may not be known by all candidates.

Here is an example of an item that contains inappropriate language. This item classifies a client by diagnosis. Note, "diabetic 50-year-old male client." Instead, we would use "client with diabetes." In addition, 50-year-old is not necessary in this case. Age is not relevant.

As we know, clients can have diabetes at any age. Note here that "male" is not necessary. Diabetes is not a sex-specific diagnosis. We rely on the nurses in this panel to determine whether inclusion of a client's sex and age are necessary for the item. Note here the use of the term "flip-flops."

Flip-flops is slang and would not be included on the REx-PN exam. "All things in moderation" is an idiomatic expression and would not be included on the REx-PN exam. Now, let's talk about underlying assumptions and stereotyping. Our panelists are instructed to flag any item that implies a generalized belief about a particular group of people that represents an oversimplified opinion, prejudice attitude, or uncritical judgment.

This is an example of an item that contains an underlying assumption or stereotype. Here, use of the term "same-sex" is not necessary and reflects stereotyping. Next, let's discuss ethnocentrism.

Our sensitivity panelists are instructed to flag any item that supports the attitude that a particular group, ethnicity, or nationality is superior to others. It cannot be assumed that all REx-PN candidates subscribe to Canadian norms. This is an example of an item that reflects ethnocentrism.

Option number one implies that this is a cultural norm that is appropriate across all cultures. This item reflects content that would be flagged and removed from the REx-PN exam. Now, let's discuss inflammatory language or content. Finally, our sensitivity panelists are instructed to flag any item that is patronizing, displaying or indicative of an offensively condescending manner such as calling a person honey or sweetie.

Condescending language is language that implies that one is conscious of descending from a superior position, rank, or dignity. Offensive language is language that is disapproving to the moral sense, good taste, or is insulting, such as a cruel joke. Finally, we ask our panelists to flag any item that contains inflammatory content. For example, statements that support or refute one's political opinion.

This is an example of an item that contains inflammatory language or content. The statement, "Don't worry sweetie," is patronizing and could be considered condescending. Now, let's talk about how the sensitivity panel works. Panelists are asked to individually review a group of items.

The panelist reviews one item at a time and decides if the item is approved or not approved based on the four criteria just reviewed. If approved, the panelist will move to the next item. If not approved, the panelist is asked to provide a comment specifying which sensitivity review criteria has been breached.

After all panelists have reviewed the items, the items that are not approved, the flagged items, are reviewed by the group. The group will discuss each flagged item, and the group will reach consensus about whether the item is approved or not approved. The approved items are used for pre-testing. The flagged items are returned to the beginning of the item development process for revision. Now, I would like to introduce Dr.

Hong, who will talk about differential item functioning.

- [Dr. Hong] Thank you, Latrice. I'm Dr. Hong. I will talk about differential item functioning, DIF. Sounds very psychometric. Make sure you have enough coffee around you.

What does DIF mean? We know each item should measure only the knowledge, ability, or skills in nursing. If this is the case, then candidates with comparable nursing knowledge will perform similarly on exam items, regardless of their gender, ethnicity, or other irrelevant factors.

But that is not always the case. DIF means that an item behaves differently for various examinee groups of equivalent ability. The keyword here is equivalent ability. If nursing abilities of two groups are not

the same, of course, they will perform differently on items. But the problem here is two groups with similar nursing ability from other items, but they perform differently on this one particular item.

Then we need to investigate this item. When talking about DIF, we always compare two groups, reference group versus focal group. Reference group is usually traditional demographic group such as females or white. Reference group needs to compare to the focal group.

Focal group can be males or non-white. Remember, DIF is always comparing two groups, reference group versus focal group. Here is a graphic representation of DIF. DIF is always comparing two groups of people.

In this example, we compare females and males. Female is reference group, and male is focal group. On the horizontal axis, we have item difficulty for female candidates from -4 to +4, the easiest to the hardest. On the vertical axis, we have item difficulty for male.

The same scale, from the easiest to the hardest. Think about it. If an item is the same difficulty level for both men and women, then the blue dots, which are items, should all fall on a diagonal line, which means the item men find equally easy or equally hard as women.

From this graph, you can see most items fall on the diagonal line, which means they are equal difficulty for both groups. But every once in a while, you will see items not on the diagonal line, item easier for one group than another group. For example, the items on the top left.

If you draw a line vertically down, you will see, for women, this item difficulty is about -1.5, which is pretty easy. If you draw a line horizontally left, you will see, for men, the item difficulty is about 2.5. That is a lot harder than -1.5.

So you know this item is favoring women. Women find this item easier, and men find this item difficult, even though these two groups have a similar nursing ability. What do we do? We pull out this item, send this item to DIF panel for review. If there is something about the content that makes the item behave this way, sometimes it is, sometimes it isn't.

If an item shows a statistically significant amount of DIF, then it is reviewed by a DIF panel to determine if there is an identifiable problem with the content, or the fancy statistics, it's just a mathematical way to do what I just told you in previous slide, picking out all the items that favor one group significantly.

If a panel determines that there is bias, the item is reviewed by Practical Nurse Exam Committee, PNEC. PNEC is REx-PN oversight group. They make all the decisions about whether an item go or stay. Bias is bad.

Difference may not be bad. Difference can be biased, but it may not be. If an item shows DIF, it could mean three things. First, although the statistics were positive for DIF, there's no evidence from reading the item itself that DIF exists.

When the DIF panel reviewed the item that, identified by statistical methods that it has DIF, the panel thinks, "There is nothing wrong with the item after reviewing it." And this is pretty common. Second, there is a genuine group difference, but nursing content requires that concept to be tested.

For example, for items about pregnancy or child care, women routinely do better than men, which makes sense. But as an entry-level male nurse, you still need to know how to deal with that. So we keep this item.

Conversely, there are some items men routinely do better than women. For example, items talking about medical equipment, how to use technology. We keep those items too because it is important for entry-level nurses to know. Whether you are men or women, which ethnicity, you still need to do the job, and it is the same job.

Therefore, even though we find a DIF item, as long as nursing content requires this concept, we still keep this item, and this item is not biased. Third, there is an extraneous feature, which is irrelevant to nursing, within the item causing the problem.

Sometimes, it is a wording issue or an unusual terminology that is not necessary for nursing practice, which causes DIF. In this case, the item is biased and needs to be revised. Here are previous findings on nursing content that has shown DIF in the past. If you look at the list, it makes sense, right?

For these concepts, you are expecting DIF items. But again, these concepts are required by nursing practice. Even though they have DIF, they are not biased and we still keep these items. Let's summarize DIF review process. First, DIF panel will read the item and indicate the groups showing DIF such as female versus male.

The second step is to form hypotheses about why the item may have shown DIF for or against a group. The panel needs to find out why this item is showing DIF. If the reason belongs to one of the first two situations, that's fine.

If there is a bias, they need to do the third step, vote to refer the item to the PNEC. Again, PNEC will make a decision to keep the item or kill the item. Now, let's review the objectives.

By the end of this presentation, you will be able to, first, identify the purpose of conducting sensitivity panel review and a DIF panel review. Now you know the reason is to make sure that exam is fair to everybody. Second, understand what it means for an item with DIF.

Now you know that DIF means two groups with similar nursing knowledge but perform differently on the item. Does it necessarily mean bias? No. Because sometimes there is a genuine group difference on an item, but the item is required by nursing practice and we keep this item.

No matter which gender you are or which ethnic group you are from, you still need to do the same job. Third, describe how NCSBN conduct sensitivity panel review and DIF panel review. Now you know these two processes. Thank you very much for joining us in this session.

I hope you didn't fall asleep. You can enter any questions you have in the Q&A box. We will have a live Q&A session coming up next.