

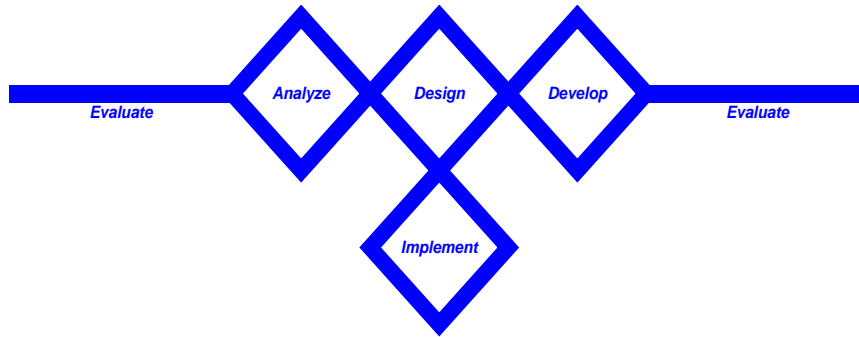
*The 2011 Working Dog Conference  
Defining, Developing and  
Documenting Success in Working Dogs.*

**A Practical Guide for Creating Performance Metrics  
and Valid Tests**

**William E. Lewandowski**  
1353 Sterling Place  
Daytona Beach, FL 32114  
Office: +1-386-868-3703  
Cell: +1-304-283-3487  
Fax: +1-866-642-8899  
email: [bill@welewandowski.com](mailto:bill@welewandowski.com)  
[www.welewandowski.com](http://www.welewandowski.com)

# Disclosure of Affiliations

William E. Lewandowski Consulting



UNIVERSITY OF  
LIVERPOOL

**Simbionix**<sup>TM</sup>  
Simulation to Advance Clinical Performance 

## Some key questions we will explore today

- **Why are you testing?**
- **What are you trying to measure or assess, and how will you measure or assess it?**
- **How will you create, conduct and administer the test?**
- **Does the test really measure what it is suppose to measure?**
- **Is the test perceived as being relevant and fair by all stakeholders?**
- **How will the results of the test be used?**

## Examining Tests



## What is a test?

### verb

• *Take measures to check the quality, performance, or reliability of (something), especially before putting it into widespread use or practice.*

• *Judge or measure (someone's proficiency or knowledge) by means of a test.*

*Reveal the strengths or capabilities of (someone or something) by putting them under strain.*

- **Oxford Dictionaries**

## What is a test?

### noun

- *A procedure intended to establish the quality, performance, or reliability of something, especially before it is taken into widespread use.*
- *An event or situation that reveals the strength or quality of someone or something by putting them under strain.*
- *A means of establishing whether an action, item, or situation is an instance of a specified quality, especially one held to be undesirable.*

- Oxford Dictionaries

## What is the purpose of a test?

<http://www.youtube.com/watch?v=OXRi28W-ENY&feature=related>

## What is the purpose of a test?

- **In regards to performance of a task, a test is generally expected to measure the ability to correctly perform a task, under real-world conditions, to a particular standard.**
- **This combination of the task, conditions and standard is called a Performance Objective. The term Performance Metric, though often used interchangeably, usually refers to an organization's ability to perform.**
- **Even the most simplest of tasks usually requires some level of mastery of a combination of skills, knowledge and attitudes.**
- **One can test to assess one or more of the constituent skills, knowledges or attitudes in a given performance objective.**

## Skills, Knowledges Attitudes (SKAs)

<b>SKA</b>	<b>Typical Verbs</b>	<b>Stimulus Examples</b>
Attitudinal	Choose	Present an opportunity to choose.
Cognitive	Discriminate, ID, Classify, Apply	Color, feel, location, texture, shape, displays
Cognitive Strategy	Originate	Novel challenge in a free-play environment.
Psycho-motor	Execute, Perform	Real-time sensory input
Verbal information	State, Recall	Menu lists, commands.

## SKAs and Performance Objectives



**A key component of creating a test is to identify the Skills, Knowledges and Attitudes (SKAs) required to master the target performance objective, so that you can then design a test to most efficiently and directly assess it.**

## Validity – What does the test really mean?

*•Validity is the extent to which a test measures what it claims to measure. It is vital for a test to be valid in order for the results to be accurately applied and interpreted.*

- About.com

**•If the results of a test are not trusted and believed by both stakeholders and independent 3<sup>rd</sup> parties, it has little value.**

**•Generally speaking, the only way to prove the validity of a test is through documentation.**

## **Does a hard test make for a valid test?**

[http://www.youtube.com/watch?v=j02\\_1\\_9xdxY](http://www.youtube.com/watch?v=j02_1_9xdxY)

## Some types of Validity

- **Face Validity** - the test appears to test takers, and stake holders, to measure what it is supposed to measure. (i.e. test takers feel that the test really measures their capabilities)
- **Construct Validity** - the test scores are a measurement of what the test is trying to measure. (i.e. an intelligence test written in English, but given to non-English speakers, would not have construct validity)
- **Content Validity** - when a group of recognized content subject matter experts has verified that the content of the test is accurate, and that the test measures what it is supposed to measure.

## Some types of Validity

- **Concurrent Validity** - the test demonstrates the ability to correctly classify experts and non-experts.
- **Predictive Validity** - the test demonstrates the ability to accurately predict future competence.
- **In the US, Content Validity is considered to be the most important form of validity in a legal defense.**
- **In the US, if a test is used in conjunction with employment, then Construct Validity is also very important.**

## Certification

- **Certification is a formal validation of knowledge or skill based on performance on a qualifying test.**
- **It is usually a means for measuring a person's competence against a given standard so as to measure an individual's performance in terms of specific skills.**

## Types of Certification Testing

- Real World

- Represents the highest level of certification testing and requires access to, and testing on, equipment or situations configured to match job characteristics.
- Accomplished in “real time.”
- Requires the testing organization to muster the appropriate resources and time to conduct the tests.
- May place either the examinee, or others, at risk.
- An example is a pilot's first solo flight.

## Types of Certification Testing

- High-Fidelity Simulation

- Represents the next, lower level of fidelity. It is a compromise in fidelity from Real World testing.
- Under the appropriate conditions, it can accurately assess proficiency across the entire spectrum of job duties, particularly those requiring psychomotor, problem solving and creative thinking skills, and at lower costs and risks than Real World testing.
- Examples of this type include a pilot using a flight simulator or a “full dress” field exercise.

## Types of Certification Testing

- Scenarios

- These are tests that require the examinee to select a best option at levels of performance directly above memorization such as: using rules, problem solving, and creative thinking.
- Scenarios are a further departure from the fidelity and cost of a Real World test, and they represent the lowest level of certification that directly assesses an individual's ability to perform a job.

## Types of Certification Testing

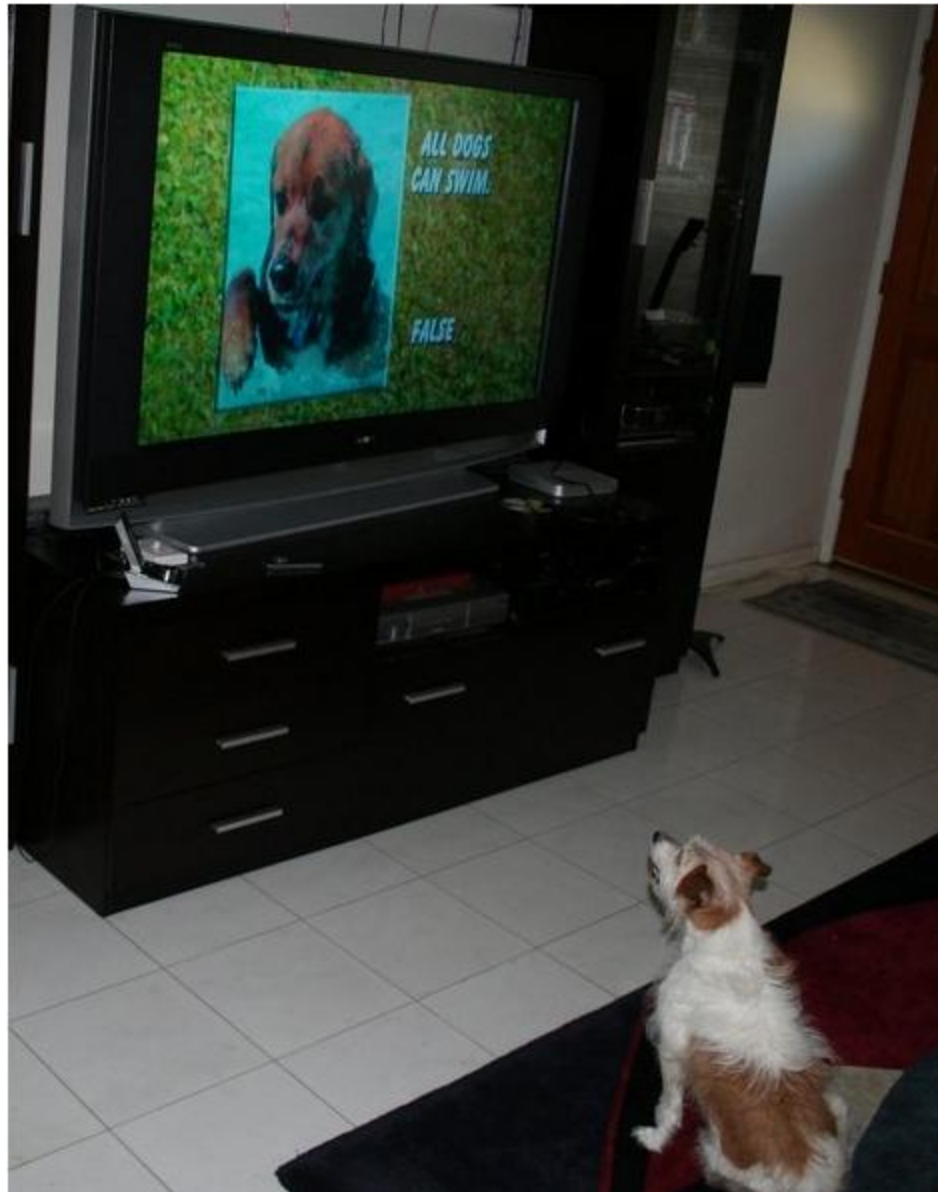
- Memorization tests

- These tests represent the lowest fidelity of certification testing, however, they are generally the cheapest to develop and administer.
- Few jobs rely solely on the recall of information, therefore demonstrating the ability to correctly recall a fact, definition or rule does not necessarily demonstrate competence in being able to correctly apply that information on the job.

## An example test

[http://www.youtube.com/watch?v=DZTkflzRJ\\_0&feature=related](http://www.youtube.com/watch?v=DZTkflzRJ_0&feature=related)

## Steps for a creating a valid test



## Steps for a creating a valid test

- **Create and maintain Plan Documentation**
  - **Plan to document the test development process.**
  - **Record each step in the process and document all key decisions.**
  - **Document and record the administration of each test and its results.**

## Steps for a creating a valid test

- Analyze the job content

- Break down each job into its constituent components and create a Performance Objective for each.
- Highlight any Performance Objectives they may cause injury, property damage, or other type of unacceptable outcome if incorrectly performed. Special care must be taken that these Objectives are attended to.
- Rank order the remaining Performance Objectives based on each one's relative criticality.
- Have subject matter experts validate this Performance Objective analysis.

## Steps for a creating a valid test

- Create and validate test items
  - Identify the SKAs required for each Performance Objective.
  - Develop a testing strategy based on the relative ranking and criticality for each Performance Objective, its SKAs, as well as any organizational requirements and constraints and any legal or regulatory requirements.
  - Pay special attention to the conditions stated in each Performance Objective.
  - Create the test instruments and items.
  - Have subject matter experts validate the test strategy, test instruments and test items.

## Steps for a creating a valid test

- Conduct an initial Pilot Test

- Identify a group of representative (12 to 15) test takers that match the demographics and skills of the target audience.
- Have them take the test under the same conditions in which you expect to administer the final test.
- Attend to any test items with high failure rates or those which generate user complaints. Revise the test as required.
- If you are planning a test with complex logistics or coordination, this a good opportunity to “flush out” the details and identify issues.

## Steps for a creating a valid test

### • Establish Cut-Off Score

- This may, or may not, be a necessary step depending on your circumstances. It can be difficult to do without training.
- Have group of experienced master performers and a group of randomly chosen (target group) examinees take the test.
- Compare the performance of the two groups. The initial Cut-Off score is that point where the lowest performing master and highest performing novice meet. Adjust according to organizational needs.
- Consider adding additional weight to critical, or high risk items.

## Steps for a creating a valid test

- **Reporting Test Scores and Improving the test**
  - For certification tests, consider only reporting “pass/fail” and using the raw scores only for statistical purposes.
  - Test scores, along with error rates for individual items, should be periodically monitored for indications of issues with specific test items or more general issues.
  - Attend to any test items with high failure rates or those which generate user complaints. Revise the test as required.
  - If you are required to conduct an advanced validation study, seek the assistance of someone knowledgeable

# Summary



## Summary

- **Creating a valid tests starts with analysis of the task to be assessed within a real world context.**
- **This analysis creates a series of Performance Objectives which is the task that must be performed under specific conditions to a specified standard.**
- **Certification is a formal validation of knowledge or skill.**
- **To be valid, a test must be shown to consistently measure what it claims to measure.**
- **Using a systematic, structured approach for creating a test will likely lead to a valid, defensible test that will provide meaningful results.**

# Questions?



## References

**"The Systematic Design of Instruction, 6th Edition" by Walter Dick, Lou Carey and James O. Carey. ISBN: 0-205-41274-2**

**"Criterion-Referenced Test Development - Technical and Legal Guidelines for Corporate Training and Certification, 2nd Edition" by Sharon Shrock and William Coscarelli with a legal analysis by Patricia Eyres. ISBN: 1-890289-09-4.**

**"The Legal Handbook for Trainers, Speakers, and Consultants: The Essential Guide to Keeping Your Company and Your Clients Out of Court" by Patricia Eyres. ISBN: 978-0070635128.**

## Multimedia Sources

**Slide 4:** <http://www.sunprairie.k12.wi.us/faculty/njgrund/>

**Slide 7:** <http://www.youtube.com/watch?v=OXRi28W-ENY&feature=related>

**Slide 10:** <http://www.vishirailgroup.com/page7.html>

**Slide 12:** [http://www.youtube.com/watch?v=j02\\_1\\_9xdxY](http://www.youtube.com/watch?v=j02_1_9xdxY)

**Slide 20:** [http://www.youtube.com/watch?v=DZTkflzRJ\\_0&feature=related](http://www.youtube.com/watch?v=DZTkflzRJ_0&feature=related)

**Slide 21:** <http://www.Doggieblogreview.com>

**Slide 28:** <http://www.funnyandjokes.com>

**Slide 30:** <http://www.DogBedX.com>