

The ERIC Thesaurus – A Key to Finding Resources in ERIC

Speaker Notes and Q&A

Presentation

Slide 2

Good afternoon everyone, I am Erin Pollard and I am the Project Officer for ERIC, in the U.S. Department of Education. Today we are going to discuss the ERIC Thesaurus. We are going to discuss how it's used in indexing ERIC records, how it's maintained and updated, and how it can help you find relevant resources in ERIC.

Slide 3

Today, three of us will be speaking to you. As I mentioned earlier, I am Erin Pollard and I am the Project Officer for ERIC. I oversee the program from the government's perspective and set policies for the program. I am going to provide some framing of why ERIC has a thesaurus and why it matters. Then Lydia Hellrich-Dawson and Judy Beck, who use the ERIC Thesaurus extensively in their daily work, will provide a deeper look.

Slide 4

Today's agenda is as follows. I am going to talk about why we have a Thesaurus and what it means for ERIC users. Then I am going to turn the presentation over to Lydia to give some background and history, and talk about the process used for maintaining and updating the Thesaurus. Judy will discuss the role of the Thesaurus in resource discovery, including using it for indexing and as a tool for searching. We will then have a wrap up and time for questions during the last 15 minutes of the webinar, but feel free to use the chat box if you have a question at any point during the webinar.

Slide 5

So to begin– why are we having this webinar? And why does ERIC need a thesaurus?

Since we are deeply involved in updating ERIC Thesaurus this year, we thought this was a perfect time to give this webinar.

We realize that the vast majority of people who search on the ERIC website don't directly use the Thesaurus to help focus their searches. The ERIC search engine uses intuitive technologies, and will bring back relevant records without having to directly use the Thesaurus. However, even if you are not searching the Thesaurus, it is being used to help with the retrieval of relevant records. This is particularly important in a broad collection like ERIC that has content spanning more than 50 years.

We also realize that even among researchers who do use the Thesaurus, there is often little understanding of how the Thesaurus is used in ERIC, and the kind of work that goes into making records discoverable despite wide variances in the vocabulary in the indexed literature. We would like to provide some background on the usefulness of the Thesaurus, as well as give you an overview of the work and care that goes into indexing records and maintaining it.

In essence, the ERIC Thesaurus is a discovery tool. It helps structure the subject matter represented in the ERIC collection to help users find relevant records. ERIC made the decision 50 years ago to incorporate a thesaurus with its database of resources. The founders of ERIC recognized that without a thesaurus, it would soon become extremely difficult to find the information you want in this rapidly growing collection of education-related reports and other resources. The Thesaurus continues to perform this necessary function today.

I'm now going to turn the presentation over to Lydia to talk about the background and history of the Thesaurus, how we maintain it, and the 2015 update.

Slide 6

Thanks, Erin.

In this first section, I will introduce you to the ERIC Thesaurus and provide a brief overview of its history. Then we will go over the different types of Thesaurus terms and what information accompanies each term.

Slide 7

The *Thesaurus of ERIC Descriptors*, or the ERIC Thesaurus for short, is a list of carefully researched and selected words and phrases that are used to describe and access ERIC documents by subject.

- The Thesaurus represents the subjects covered by ERIC. From the Thesaurus, we can tell what subjects are covered by the documents in the ERIC collection.
- The Thesaurus organizes ERIC documents by subject. Each term in the Thesaurus represents a subject covered by ERIC and, therefore, can be used through the process of indexing to organize the ERIC collection by subject.
- The Thesaurus was designed for search. As we will discuss in a moment, the Thesaurus was designed as a tool for information retrieval. It has always been used as a way to access ERIC documents by subject.

Slide 8

The ERIC Thesaurus is almost as old as ERIC itself. Preliminary work to establish a thesaurus of education terms for ERIC indexing was conducted in 1965, making it one of the earliest controlled vocabularies developed for an information retrieval system. The first version of the Thesaurus was released in 1966. There were 14 print editions of the Thesaurus released between 1966 and 2001.

The ERIC Thesaurus is recognized as an authoritative vocabulary tool in the field of education and, as such, has served as a reference for the construction and development of other thesauri, including the *Australian Thesaurus of Education Descriptors* and the *British Education Thesaurus*.

In 2004, the Thesaurus went fully digital and was made available to search and browse on the ERIC website.

Since then, the Thesaurus has been regularly updated with new versions and released directly to the ERIC website.

Slide 9

Today, the Thesaurus is accessible from the ERIC home page. The Thesaurus is also available for download by the general public (including other ERIC information providers) in the Download section of the ERIC website.

Later, we will talk about how we maintain the Thesaurus and the upcoming Thesaurus update.

First, let's talk about the terms that make up the Thesaurus.

Slide 10

The Thesaurus contains two main types of terms:

- First, we have Descriptors. A Descriptor is a word or phrase that represents a subject in the field of education. Descriptors are the terms used to index ERIC records and they appear on every record in the ERIC collection. Only one Descriptor, or sometimes two, are used to represent a single topic. This is done so that all records on a topic, regardless of the language used by the authors, will be gathered together under only one or two Descriptors. Since we can only have one or two Descriptors per subject, it is important to have additional terms in the Thesaurus to serve as guides to those Descriptors.
- Those guides are called Synonyms. A Synonym is a word or phrase that means the same thing, or nearly the same thing, as a Descriptor. While they may not seem as important as Descriptors since they are not used in indexing, Synonyms are incredibly useful as we can have multiple Synonyms to serve as guides to one Descriptor. Synonyms can represent popular abbreviations, differences in spelling, and user-preferred terms.

Slide 11

Each term in the Thesaurus, whether it is a Descriptor or Synonym, has information associated with it that explains how the term is used in ERIC, clarifies its relationships to other concepts within the Thesaurus, and provides some historical information. We call this a term entry.

A Descriptor term entry has a lot of useful information:

- In every term entry, you will see the term itself, in this case "Adult Education."
- Descriptors frequently have scope notes, which provide a description of how the concept is defined within the ERIC literature. It may also provide references to other terms within the Thesaurus, differentiate it from similar terms in the Thesaurus, or share historical information about its use in indexing.
- Each Descriptor is assigned to a Category. The categories group conceptually related Thesaurus terms together as a tool to help users locate useful Descriptors. We will talk more about using the Categories later.
- Descriptors may be related to other terms in the Thesaurus.

It may have Broader Terms, indicating that the Descriptor is an example of or an instance of a more comprehensive term. Broader terms are less precise than the Descriptor.

The Descriptor may have Narrower Terms, indicating that the Descriptor is a more comprehensive term than others in the Thesaurus. Narrower terms are more precise than the Descriptor.

The Descriptor may also have Related Terms. These are other terms in the Thesaurus that may be associated with the subject represented by the Descriptor and may be useful in selecting terms for indexing or searching.

- Under the heading Use this term instead of, the Descriptor's Synonyms are listed.
- The date that the Descriptor was added to ERIC is also provided.

Please note that not all Descriptors have all of the elements. A Descriptor may not have a scope note, or any related terms, or any Synonyms. That is not a mistake – that just means that when the Descriptor was added to the Thesaurus, those elements were not deemed necessary or appropriate.

Slide 12

A Synonym's entry may not have as many elements as a Descriptor entry, but you will always find the most critical element: the Descriptor that is used instead of that Synonym in indexing.

In this example, the Synonym "College Community Relationship" directs you to use the Descriptor "School Community Relationship."

Slide 13

The Thesaurus is not static or set in stone. As I mentioned earlier, over the years the Thesaurus has changed, with revisions and updates released periodically throughout its history.

In this next section, I will talk about why and how we maintain the ERIC Thesaurus. First, we will review the reasons it is important to update the Thesaurus. Next, we will go through ERIC's Thesaurus update process. Finally, we will take a "sneak peek" at the upcoming 2015 Thesaurus update.

Slide 14

Why update the Thesaurus? There are three main reasons why it is important to regularly update the ERIC Thesaurus: usability, currency, and relevancy.

Slide 15

The first reason is to ensure the usability of the Thesaurus – to make sure you can find the terms you want so you can find the content you want.

Additions to the Thesaurus, whether they are Synonyms or Descriptors, enhance access to ERIC content.

This is why Synonyms are particularly critical additions to the Thesaurus. As I mentioned earlier, we can have many different Synonyms leading to one Descriptor. By adding additional Synonyms, we can increase the chances that users will find a Descriptor for their topic of interest in the Thesaurus

Updating the Thesaurus to enhance its usability is critical to preserving ERIC's overall usefulness to its users.

Slide 16

Another reason to update the Thesaurus is to keep it current.

Language is constantly evolving and the language of education is no exception. As I mentioned earlier, the terms used in education research change over time. They may reflect changes in attitudes and interests. Once-popular topics may fall by the wayside. Some topics become obsolete as research in the field advances.

Keeping the Thesaurus current is critical to ensuring that the Thesaurus reflects the active vocabulary of education and that ERIC can help provide access to its content using the most up-to-date terminology.

Slide 17

Finally, it is important to update the Thesaurus to keep it relevant.

New fields of interest and areas of research need to be incorporated. The boom in technology use in education, for instance, has required, and continues to require, a number of new terms.

Education research is also quite interdisciplinary causing terms in other fields to bleed into the vocabulary of education.

A relevant Thesaurus ensures that ERIC continues to provide users with easy access to content in emerging topic areas.

Slide 18

How does ERIC go about updating the Thesaurus?

There are three main steps in the process: Term Identification and Selection, Proposal Development, and Proposal Review.

Slide 19

First, we need to identify terms to be added and terms to be changed. Our primary sources of information are ERIC indexers and users:

- ERIC indexers see ERIC content every day and note where there are gaps or changes in the vocabulary.
- ERIC users bring perspectives from research as well as practice and provide subject area expertise.

Indexers and users may request:

- New Descriptors
- New Synonyms
- And changes to existing terms. Changes to existing terms include adding or revising scope notes, adding or revising relationships, or transitioning terms from Synonym to Descriptor status, or vice versa.

Slide 20

Once we have identified potential additions and changes to the Thesaurus, we need to select a group for further development.

All indexers and user requests are reviewed and the final selection is made based on criteria established in the ERIC User Guide. For each new term request, we ask:

- Does it qualify as a Thesaurus term? Thesaurus terms must be nouns or noun phrases that represent a subject in the field of education. Proper nouns are not usually included, although they may be if they represent a significant concept in education research.
- Is It Useful? Terms added to the ERIC Thesaurus must be useful for indexing and retrieval. To determine if a term is useful, we consider:
 - Does it duplicate any existing Thesaurus terms?
 - Does it exhibit staying power, showing regular usage in the education literature over time?
 - Is it an emerging term that is replacing an older term in the education literature?
 - Is it a Synonym of an existing Thesaurus term that would aid in retrieval?

We also ask:

- Is there literary warrant? Literary warrant means the term actually appears in ERIC content. A term must appear in the ERIC literature before it can be added to the Thesaurus. Terms with a high level of literary warrant – 1000 hits or more – are considered to be priorities.
- Is there a need? If we've received multiple requests for the same addition or change, we also consider those requests to be a priority.

Although we can't select every request for further development, we never get rid of any requests. All requests are kept and examined when a Thesaurus update is scheduled.

Slide 21

Once a term is selected for further review, a proposal for how the term should appear in the Thesaurus must be developed.

First, research is conducted on the term. Definitions in dictionaries, encyclopedias, other thesauri, and ERIC documents are examined and documented. This helps to establish a working definition of the term that can be used to construct a scope note and clarify relationships to other terms in the Thesaurus. The working definition may also help determine if a term requires a scope note to distinguish its use in the field of education from other uses or to differentiate it from another term in the Thesaurus.

An example of the latter in the Thesaurus is the term "Performance Based Assessment" which includes a note to distinguish it from "Performance Tests."

Slide 22

Term usage in ERIC is examined to identify any differences in how a term is defined or used within the ERIC literature.

This helps determine if more than one term is required to address multiple uses. An example from the Thesaurus is the concept of articulation which has two different definitions in ERIC. Two terms were

required to cover both definitions, so two terms were added, each one with an indication of its definition provided in parentheses. The categories and the scope note for “Articulation (Education)” further distinguish these terms from one another.

Examining term usage also helps to determine if a term is not used consistently, indicating that more time may be needed for the concept to solidify before it can be added to the Thesaurus.

Slide 23

After we complete the research, we use that information to construct the term entry. This includes:

- Selecting the final form the term name will take
- Writing a scope note if necessary
- Establishing the appropriate relationships
- Selecting a category

All of this work is completed by experienced professional taxonomists. We check each other’s work, discuss any issues, and must come to a unanimous agreement on all aspects of a term’s entry before it can proceed to the next step.

Slide 24

Once the taxonomists have completed their work, the proposals are then reviewed by the ERIC indexers and quality assurance staff. They thoroughly review the proposals, checking the research and the proposed term entry.

Having two layers of review as well as multiple reviewers with different expertise helps ensure that the changes we are making to the Thesaurus are accurate, complete, and that they meet the needs of indexers and users.

At this stage, some changes to the term entry may be made if there are concerns about the entry as proposed. Once we have resolved any issues and achieved a majority of approvals for a proposal, it is then considered ready for official inclusion into the Thesaurus.

Slide 25

The work to update the Thesaurus began in February, starting with the term identification and selection process. As part of that first step, we invited the user community to provide input into the update by contributing their suggestions for new terms and changes to existing terms. A number of those suggestions were selected for further development and are anticipated to be part of the completed update.

The Proposal Development stage was completed in June. Now, we are in the Proposal Review stage. A final list of recommended additions and changes to the Thesaurus will be sent to the Department later this month. The Thesaurus update is slated for release to the web site later this year.

Slide 26

The 2015 update will include nearly 20 new Descriptors.

Proposed new Descriptors include:

- Common Core State Standards
- Low Income Students
- Value Added Models

Slide 27

The update will also contain nearly 30 new Synonyms.

Some of these are being entered as Synonyms for new Descriptors. Some are being added as Synonyms for existing Descriptors.

Proposed new Synonyms include:

- Differentiation which is a new Synonym of Individualized Instruction
- MOOCs (an abbreviation of Massive Open Online Courses) which will be a Synonym for Online Courses

Slide 28

Finally, there will be approximately 17 changes made to existing terms.

One of the most important changes we are making is to switch from using the term “Mental Retardation” to the term “Intellectual Disability.” We’ve been tracking the transition in the literature for several years. “Intellectual Disability” was added as a Synonym in 2008 to accommodate the growing use of the term and in anticipation of eventually making it the preferred term. Once the update is complete, “Intellectual Disability” will be a Descriptor and “Mental Retardation” will be one of its Synonyms.

This chart shows the number of records in ERIC containing “mental retardation” or “intellectual disability” in the title or abstract from 1966 to 2012. As you can see there was a clear transition from “mental retardation” as the preferred term to “intellectual disability” as the preferred term.

Now that we’ve covered the nuts and bolts of the Thesaurus, Judy Beck will talk about how the Thesaurus is used to enable resource discovery in ERIC.

Slide 29

Thanks Lydia.

In this section, we will talk about how the Thesaurus is used in indexing ERIC records and how the Thesaurus then aids in searching.

Slide 30

Descriptors from the Thesaurus are assigned to every ERIC document.

If you have seen an ERIC record, you have seen a list of ERIC Descriptors—every record (with few exceptions for early records) has at a minimum 3 Descriptors assigned to it.

That minimum really only applies to records where no full text and the skimpiest of abstracts is available. Otherwise, ERIC records contain the number of terms necessary to adequately describe the content – no more and no less.

Slide 31

Descriptors reflect the content of the document regardless of the language used by the author.

The field of education in all its facets contains what I call “squishy” terminology. As Lydia mentioned early, terms and phrases change meaning over time, cross paths between older and newer phrases, come and go with the latest thing, and one word can have different meanings in different contexts.

Only one or two terms are used to index a particular subject. That means that all documents on that subject will be indexed with the same term or terms.

For example, in the record pictured, the author uses the term “primary school students.” We know that primary school students are the same as elementary school students, so the record has been indexed with the Descriptor “Elementary School Students.”

Slide 32

Descriptors reflect the specificity of the document (to the greatest extent possible).

For example, if a document is about high school students, it will be indexed with the Descriptor “High School Students” – not a broader term like “Secondary School Students” or just plain “Students.”

Slide 33

Users often ask, “How do Thesaurus terms get assigned to ERIC records?”

In general, there are three ways to approach indexing: automated, machine-assisted, and manual.

Automated indexing eliminates the human factor completely and terms are assigned by a text-reading program that has been “trained” to pick up key words and concepts matched to a controlled vocabulary. This approach works best for highly precise vocabulary, such as found in medical texts.

Machine-assisted indexing uses a program to suggest terms, but an indexer reviews and makes adjustments as necessary. This approach is helpful to the degree that the terms in the text can be closely mapped to the controlled vocabulary and the program can make discernments between different meanings. Adding the human element insures that inappropriate suggested terms are eliminated and missed terms are added.

In the manual approach, all assignment of terms is done by a professional indexer—a human must look at each record and know the Thesaurus well enough to select the terms most appropriate for the content of the text.

ERIC uses a combination of manual and machine-assisted with professionally trained indexers. Most indexing is done manually. Machine-assisted indexing is used to help improve indexing consistency in cases where certain indexing terms are always used together.

Slide 34

Assigned ERIC Descriptors reflect the subject of the document in hand—whether it’s a journal article, report, conference paper, or any other accepted form of content.

ERIC Indexers use all available information to assign Descriptors to each document. In some cases, all that is available to an indexer is the title of the document and the abstract.

Titles can be helpful and descriptive—or in some cases are clever but totally not helpful for indexers.

Abstracts are key elements in indexing because they describe the content of the document. Abstracts are either provided by the author or publisher of the document or may be constructed by an ERIC information specialist.

When available, the full text of the document is also consulted for concepts not included in the title or abstract.

Slide 35

Now that you know more about the Thesaurus and how it is used in ERIC indexing, let's get searching!

There are two main ways the Thesaurus helps enable resource discovery in ERIC:

You can use Descriptors to help refine your search, or, you can use Descriptors to start your search.

Slide 36

You don't need to know specific Descriptors to benefit from using them in your search.

You can search ERIC using any criteria you like, then use Descriptors to narrow down your search results.

The ERIC search results page displays additional search criteria on the left hand side of the screen. A list of Descriptors assigned to the records in your search results appears just below the publication date.

You can select a Descriptor from that list by clicking on it and your search results will be altered to display only the records that include your keyword and are indexed with that Descriptor.

For example, a search for the keyword dropout results in 16,022 hits. On the left hand side of the screen, you can see that these records have been indexed with Descriptors including "Dropouts," "Dropout Prevention," and "Dropout Rate."

Slide 37

If we select the Descriptor "Dropout Prevention," the search is narrowed to only those records containing the word "dropout" that are also indexed with "Dropout Prevention." This has narrowed the search to 3,465 records.

You can continue to use Descriptors to narrow the search even further, until you get a set of records that meets your needs.

Slide 38

You can also use a Descriptor to start a search.

To start your search with a Descriptor, you will first need to find a Descriptor to use.

You can easily access the ERIC Thesaurus on the ERIC website. Simply click on the Thesaurus tab on the ERIC home page. From here, you can either search or browse the Thesaurus.

Slide 39

To search the Thesaurus, enter your search criteria in the box, and then click on the Search button. To include Synonyms in your Thesaurus search, click on the "Include Synonyms" check box just below the

search box. As Lydia discussed earlier, Synonyms serve as guides to Descriptors so it may be useful to include them in your search.

Slide 40

Once the search is complete, you will see a list of all Thesaurus terms that contain your search criteria. In this case, we've searched for the word "reliability" and retrieved 5 terms. Synonyms appear in italics.

From this list, you can select a term to see its complete entry.

Slide 41

Another way to access the Thesaurus is to browse it. To browse the Thesaurus, click on the Thesaurus tab on the ERIC home page, then click on the Browse Thesaurus link to the right of the search button.

Slide 42

You can browse the Thesaurus alphabetically or by category.

To browse Thesaurus terms alphabetically, simply click on one of the letters and all Thesaurus terms starting with that letter will be displayed.

Slide 43

Here we can see the display for the letter "Y" which includes the terms "Year Round Schools," "Yearbooks," "Young Children," and "Youth Programs." From this list, you can select a term to see its complete entry.

Slide 44

You can also browse the Thesaurus by Category. These Categories are broad subject areas that are used to link conceptually related terms. As we saw earlier, every Descriptor is assigned to a Category.

Slide 45

When you select a Category, you will get a list of all Descriptors that are assigned to that Category.

In this example, we've selected the Arts category which includes the Descriptors "Art Education," "Dance Education", "Graphic Arts," "Music Appreciation," and "Theater Arts." From this list, you can select a term to see its complete entry.

Slide 46

Once you've found the Descriptor you want, whether through browsing or searching, you can use it to search the ERIC collection. It is easy to do.

Just below the Category, you will see a link that says: "Search collection using this Descriptor". Simply click on the link and a search is automatically generated for you.

Slide 47

Here we can see the search results for the Descriptor "Art Education." This shows us every record in ERIC that has been indexed with the Descriptor "Art Education." From here, you can use the techniques we discussed earlier to refine your search using other Descriptors.

Remember, the Thesaurus is a discovery tool—it reflects the actual content in ERIC and helps assure that all relevant records that include a specific concept, regardless of the author's terminology, can be found.

Erin, I'd like to turn it back to you now for our wrap up.

Slide 48

So to conclude, we hope that this webinar was helpful in answering your questions about the ERIC Thesaurus and it will encourage you to try using Descriptors to focus your searches. We hope that you will spread the word about its usefulness!

Slide 49

Now we are going to take any questions that you have. As I mentioned at the beginning, we are going to answer any questions specifically on the Thesaurus at this time. If you have other ERIC questions we will answer them after all of the Thesaurus questions are answered.

If you think of a question after this presentation or think your question would be best answered offline, please email us at ERICRequests@ed.gov. We will answer your questions that way as well.

Slide 50

Thank you so much for attending this webinar. Please look for the slides and archive of this webinar to be posted in a few weeks at eric.ed.gov/?multimedia.

Questions and Answers

Following are the questions asked during the webinar and responses from the ERIC Team.

Question 1 - Can you select more than one Descriptor from the list at once?

With how the website is currently set up with the Descriptors appearing on the left side of the page, you cannot select more than one at a time. However, you can search for more than one Descriptor using the text box at the top. Or, after you have clicked a Descriptor in the list [and narrowed your results] you can then click on the next Descriptor that you want to include in your search.

Question 2 - When refining a search result with Descriptors, does the refined search show all documents with that Descriptor term which is assigned to the article, or just the documents that use the Descriptor term as an index term?

Let's say you search for "schools" and then you click on the Descriptor "Dropout Prevention." What that will do is show all search results that contain the word "schools" that are indexed with the Descriptor "Dropout Prevention." So then it's a subset of the records, it's not everything related to dropout prevention, but it's those records that meet the keyword search as well as having that Descriptor assigned to it. You could go through the Thesaurus to launch a search for all records indexed with that Descriptor.

Question 3 - Does the download of the ERIC Thesaurus include the metadata for the associated ERIC records?

For the ERIC downloads there are two sets of records you can download - two sets of XML. The first is the Thesaurus that has the structure of the Thesaurus. The second set of files are all of the metadata for ERIC. So all of that information is available, but they technically are two sets of files. And they are clearly marked when you go on the website eric.ed.gov/?download.

Question 4 - When you had "Y" Descriptors shown under the alphabetical search, why did some Descriptors show "(2004)" as part of the Descriptor phrase?

When the Thesaurus was still in print, there was a practice of including after the name of a former Descriptor (something that had been moved from Descriptor status to Synonym status) in parentheses the dates that it was used in indexing. In 2004, there was a revision of the Thesaurus and some terms that had previously been Descriptors were made Synonyms, so "2004" was put in parentheses after the term name to indicate that the term previously had been a Descriptor but in 2004 was made a Synonym.

Question 5 - Will you be introducing a search tool that would allow Boolean logic? (and/or searches)

In 2013 we redesigned the ERIC search based off of user feedback. Many of the people that were using the eric.ed.gov website were not using the traditional Boolean logic. And the searches that we have developed when you click on the Advanced Search tips we think can replicate every kind of search that users need. If there's a specific type of search that you used to be able to do, but cannot do now, please e-mail us at ERICRequests@ed.gov. We're not opposed to bringing back the traditional advanced search or Boolean search, we just need some examples of why what we have isn't working to justify why we're going to build that. Since we launched the site – almost 2 years ago this week – we've found one or two

Questions and Answers

good examples, but we're trying to find just a few more. When we skipped things that you cannot do in ERIC, we refined the search engine to try to find a way that will make the search possible and still keep with the more streamlined search approach that we find that most of our users are using.

Questions 6 - Is the Thesaurus in the EBSCO and ProQuest versions of ERIC the same as ERIC original Thesaurus?

We can't comment on what private companies are doing. There are just so many people that use ERIC, we don't know the ins and outs of how they operate and parts that are proprietary. We make our Thesaurus publicly available to EBSCO and ProQuest, as well as to every other person and company, and I would assume that they are using our ERIC Thesaurus metadata. I know that they are using our metadata for other parts [such as ERIC record updates].

Question 7 - Is ERIC still using Identifiers as well as Descriptors, and if so, what is the role of the Identifier?

We are still using Identifiers, and the Identifiers represent three different categories. They all represent proper nouns and they represent geographic locations, laws and legislation, and tests and testing. Identifiers are assigned very similarly to the Descriptors – they represent subjects in the document itself – so the geographic location where the study took place; any tests that were utilized; any laws, or court litigation that are discussed in the document itself.

So just as a little follow-up – we know that it doesn't necessarily make sense in 2015 to have all three of those things in one field. It made sense back when the field was first identified, but as we've grown and evolved we've realized that it's such a treasure trove of information. We are working to subdivide the Identifiers field. We just started work on that this week and we will be providing a lot more information long before that becomes available, but we do think that there's tremendous value in being able to search the Identifiers field based off location and what kind of tests people are using, and we're trying to find the best way to make that information available.

Question 8 - What is the best way to limit results to articles pertaining to the United States only?

That question is actually a little more complicated than it seems. Because you can look for work written in the United States, you can look at work published in the United States, or work written about Americans. Each of those queries has a different way of doing them.

As far as indexing goes (and we index using Identifiers and geographic locations), the assumption is that, unless it otherwise states, the work pertains to the United States. So we generally do not use the Identifier "United States" itself unless something is comparing the United States to another country. So if it's an article that is comparing Taiwan, for example, to the United States, we will have the Identifiers "Taiwan" and "United States" in the Identifiers field. If you really want to narrow it down to things that don't include foreign countries, you would want to eliminate the Descriptor "Foreign Countries" from your search because we use the Descriptor "Foreign Countries" whenever a document also has a foreign country Identifier. So if a work is just about Taiwan, we use the Identifier "Taiwan" and the Descriptor "Foreign Countries." You can use the minus search on the website to eliminate "Foreign Countries" as a Descriptor and that should bring back things that are only in the United States.

Questions and Answers

Question 9 - Will the term "Native Americans" replace the term "American Indians"?

That is not in the current Thesaurus update based on literary warrant and how the field is using the term. We see the term "Native Americans," many people use it, but "American Indians" is still the preferred term. When you look at reports published from the Institute of Education Sciences, both terms are still used. We want to make sure that when we update something, the field has already shifted to reflect that change.

Question 10 - Do you know when database aggregators, such as EBSCO, will update to the new Thesaurus?

EBSCO, ProQuest, and you will get access to the updated Thesaurus on the same day, at the same time, with the same message. They go to the same public download area that is also available to you. So when we make the change (it will be live on the website this fall), we will send out announcements that it's there. Then private companies and private citizens can go to the website and update their database. We make our information publically available for free, and private companies can do whatever they would like to when the information is available.

Question 11 - Will we be able to get a copy of this PowerPoint presentation?

Yes. We will be making this live at eric.ed.gov/?multimedia probably in the next two or three weeks, and we will also put up a video of the presentation so you can hear what we said. We will put up the transcript, and we will send out an email to everyone who registered for this webinar to let them know that that it is now live.

Question 12 - Once this current update is complete, will a new print edition of the Thesaurus be published?

If you're thinking about that big blue book that was last published in 2001 that will not be published. The government does not print books anymore so we won't ever have a physical hard copy. At least in the foreseeable future given the focus on the environment and unnecessary paper printing. In terms of making that information available, we have been trying to balance how people would use the print copy. In what ways would a PDF be more useful than the search function that we have, and our metadata that we have – the XML file? And we've been trying to figure out in what ways a PDF would be useful – what are people looking for? So, if you could, email us at ERICRequests@ed.gov. Let us know if you would find the PDF helpful, how you would use it, and why it would be more useful than what we currently have. We absolutely can make it available, we just don't want to take the time and effort away from other projects to make a PDF that we don't really think would be helpful. So if we understand how you'd use it we aren't opposed to bringing it back. We just can't imagine how you would use it and would love to hear more.

Questions and Answers

Question 13 - When a Descriptor is "retired" in favor of a newer Descriptor, do you go back and re-index the articles that used the old Descriptor?

No, we don't do that. As Lydia was talking about, the term "Intellectual Disability" is coming in as a Descriptor and "Mental Retardation" is being made a Synonym. That's why it is important, when you're doing a Thesaurus search, to check that box that will bring Synonyms in with your results. That would be a way to capture the older records that were indexed with the older term.

Question 14 - There's a controlled list of Descriptors, right? Is there also a controlled list of keywords or Synonyms?

To answer the last part of that question about a controlled set of keywords – a keyword is a word that people use for their search. So there is no controlled list of what people search. As far as controlled Synonyms, those all appear in the Thesaurus, as well as the Descriptors.

Question 15 - Is there a controlled list of Identifiers? Do they appear in the Thesaurus?

As of right now, there is no list of Identifiers that we've made publically available. As I mentioned earlier, we are literally just working on how to best take this treasure trove of information and make it available. We will definitely make it publicly available through the metadata when it is ready.