00:05:32 Matthew Burns - SchoolPsychedPodcast117

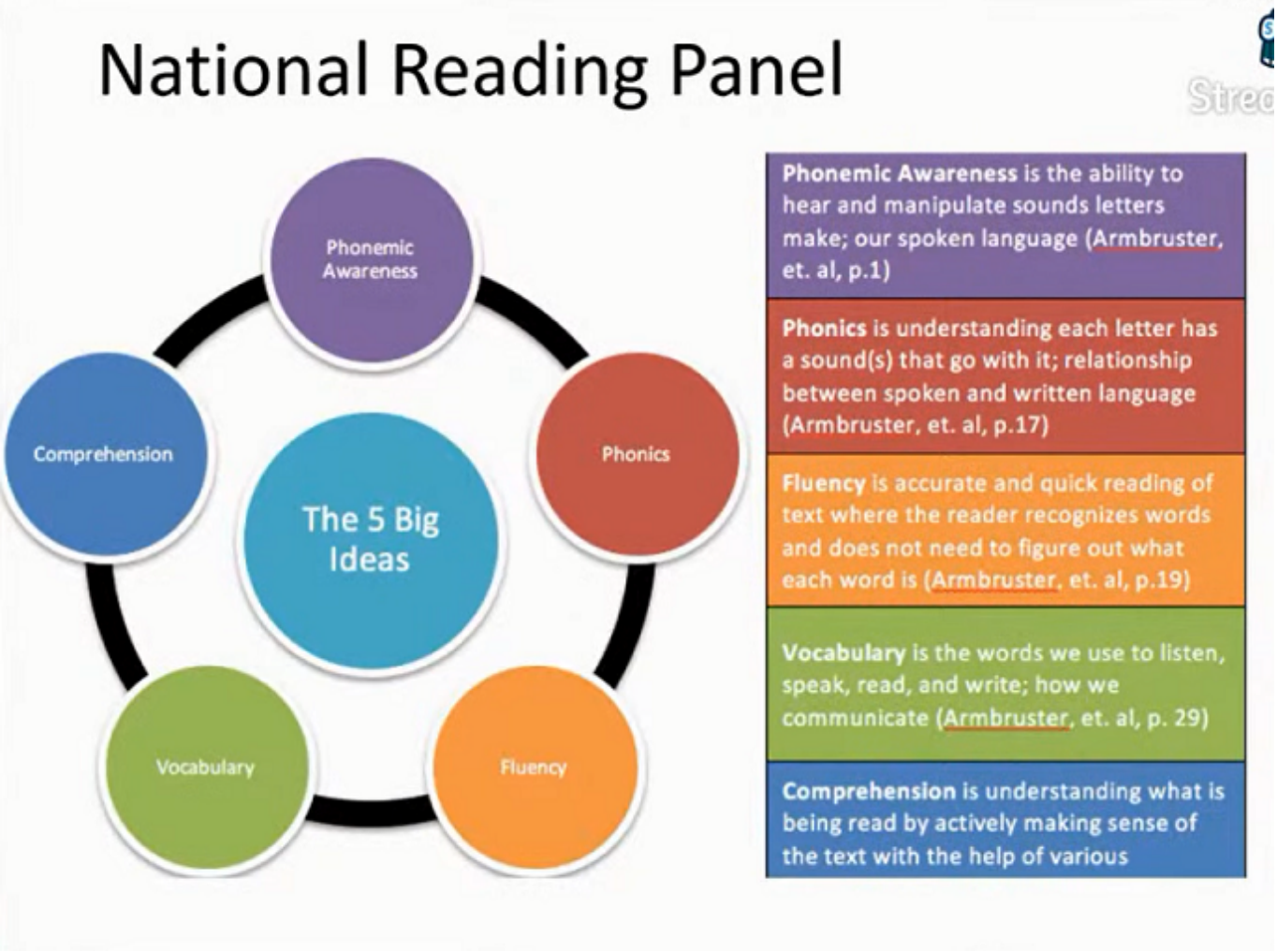
- over half of a school psychologist's time is spent on evaluations.
- 80% of learning disabilities are real
- (Burns recent study) looking at reports of school psychologists, and of the 2/3 that had recommendations , only half of them (48%) were based on data, the rest of them were generic (i.e. put the kid in front of the room)
- 2/3 of fourth graders do not score in the profficient range (National Center for Education Statistics. 2013)
    - 47% below skill level
    - 65% of English learners are below skill level
- Teacher's self-efficacy **decreases** after their first year (Clark, 2020)
    - recognizing what they *don't* know increased their efficacy

---

00:07:53 Science of Reading (SOR) as example

- SOR literature shows that effective reading instruction is very complex and individualized (Connor, 2016)

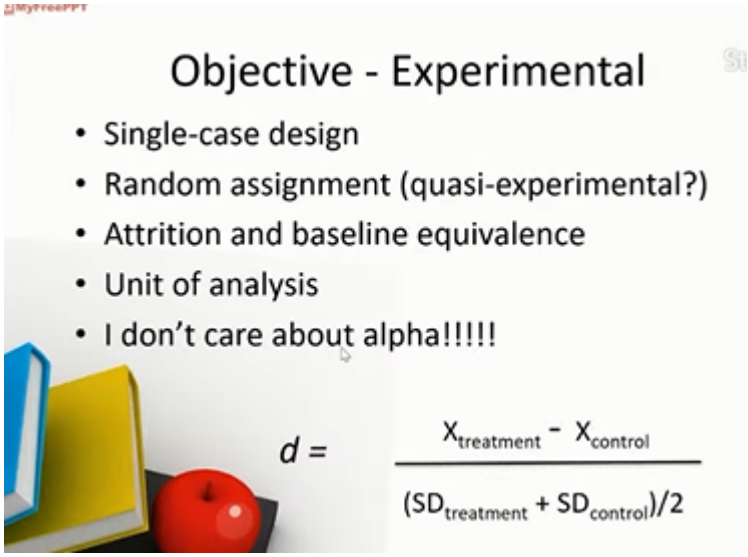
00:08:26



00:08:48 What constitutes good resources?

- Objective experimental research, the focus ought to be on something that shows **cause and effect**, and not subjective (where there is a study that outlines alearly an indepedent and a dependent variable)
- Experimental research, in other words
  - not descriptions
  - not logical arguments
  - not blogs or books

00:10:00 Example of objective experimental research



- *When consuming research, a useful skill would be the ability to recognize features of an experiemtnal design. For example, an experimental design entails that there was a random assignment of participants*
- *Another red flag for an inferior study, according to Dr. Burns, is an unusual attrition number. For example, if the study strted with N=60 and ended with N=40, this should be considered a red flag. Sinde there is such a degree of dropout from a study, this means that the two groups are no longer the same at baseline. If the baseline is not relatively similar between the two groups at the start and at the end of a study, then the baseline measure will have been compromised*
- *Another feature of a study worth discerning is whether the assignment was done by classroom as opposed to individual children, indicating that the true unit of analysis was the classroom, not the individual children. This reduces power tremenduously, and may not reach significance as a result, of course, but more importantly, understanding whether the unit of analysis was children or classrooms can be unclear sometimes in studies and it may indicate a flawed analysis.*

00:12:00

- When you're reading articles, watch for randomization, see if there'sany attrition, that is, do they have evidence to show that the two groups are the same at baseline, and the way they were analyzed by the unit. If they were assignmed by teacher, the analysis should be by teacher. If they were randomized by diet, the analysis should be by diet, or classroom to classroom etc.

00:12:27

- **Alpha vs *d***
  I have published over 200 articles and I have yet to test for the null hypothesis. Alpha is designed to only show you that the results are not due to sampling error. For example, if your Alpha is 0.06 and you sampled 100 kids what do you do? The answer is, you collect more data, since Alpha is basically the relationship between N and the size of the effect. So if you don't increase the size of the effect but you increase N, the Alpha goes down.

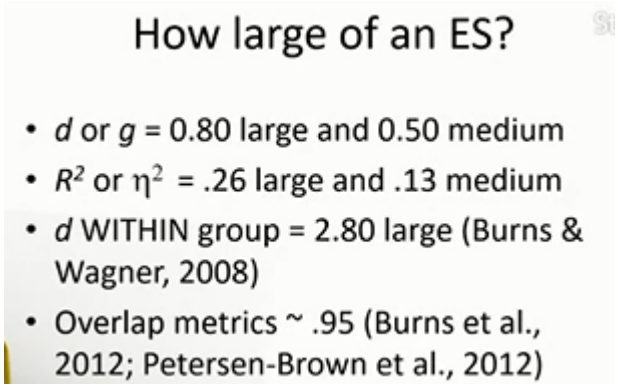
So you look at d instead, Hedges' g is even better.

This, here is not the formula for Cohen's d but if you are in the middle of reading a research paper and need a quick interpretation of the results use this easy way to do it:

$$d = \frac{X_{treatment} - X_{control}}{(SD_{treatment} + SD_{control})/2}$$

In other words, take the difference between the mean of the treatment and the control groups and divide that by the average of the standard deviations. That is not d necessarily, but when you're reading an article, if you do just this quick calculation it will be really close to the correct d unless the variances are really off, which they usually are not. The difference, in most cases will be, for example 2.6 for the this quick and easy shortcut, to the d=2.61 for the actual d, properly calculated, which is really close. This is useful in the moment, so you get a sense of the numbers you are looking at, in case the effect size is not reported, to give you a good sense of the type of effect.

00:14:52 How large is the ES?

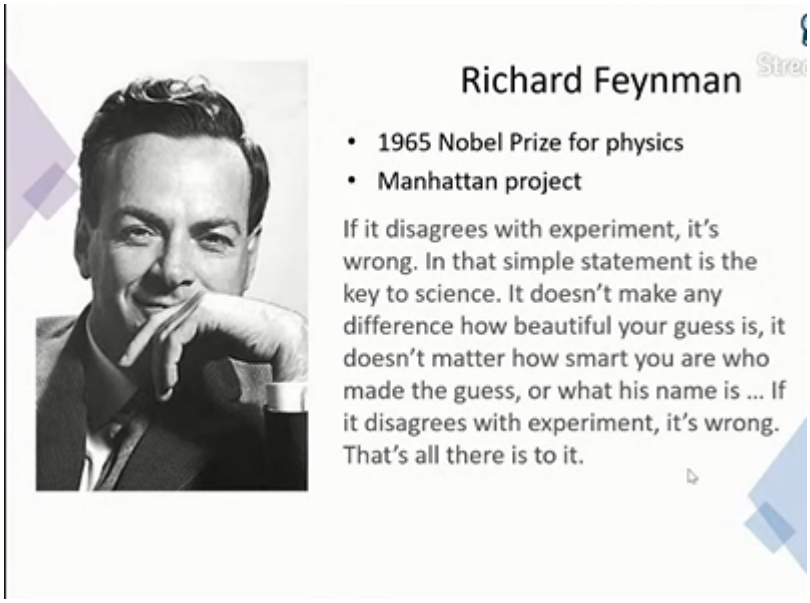


- think of Heta squared as a percent variance accounted for by the independent variable, which is an oversimplification, but for consumption it would be adequate
- Wagner, a researcher sums it best by saying: "School Psychologists need to bethe researchers in schools, htey should be the ones in the schools who can conduct research, but they should also be the ones who can consume and synthesize research." So if a principal says "I think I'm going to start a responsibility thinking

group, what does the research say?", and the school psychologists should be the ones that consume the research and talk to the principal about what the research says.

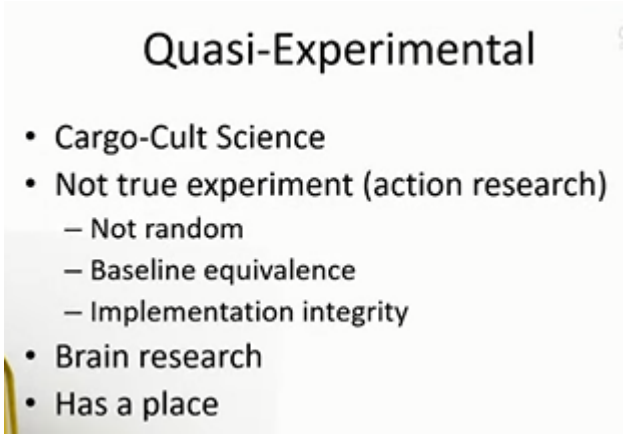
- again don't get ccaught up on these numbers, for example. a d=0.7is quite good, sometimes an R^2 is quite good.
- but generally these are acceptable ranges
- but beware of things lie the following:
    - a study does a pretest and a post-test and reports a d
    - then they report that the d = 0.8 and that it was a large effect
    - the d=.80 is interpreted as a large effect if the experiment was a between group design
    - if the experiment was a within group design, the effect size is being inflated dramatically. The average d for a large effect in within group designs needs to be 2.80. For a within group design, you need to see something lose to 3 .0 standard deviations in effect size for it to be considered a large effect.
    - Overlaping Metrics in single case design need a 95% effect size in order to consider it a large effect size.

00:18:11 On Richard Feynman



- you may want to check out "Mr. Feynman surely you're joking" and "The pleasure of finding things out".
- Feynman quote urges us to take a guess and one may think of it as a hypothesis. If it disagrees with the experiment - he says it is wrong, it's that simple.

00:18:11 Quasi Experimental

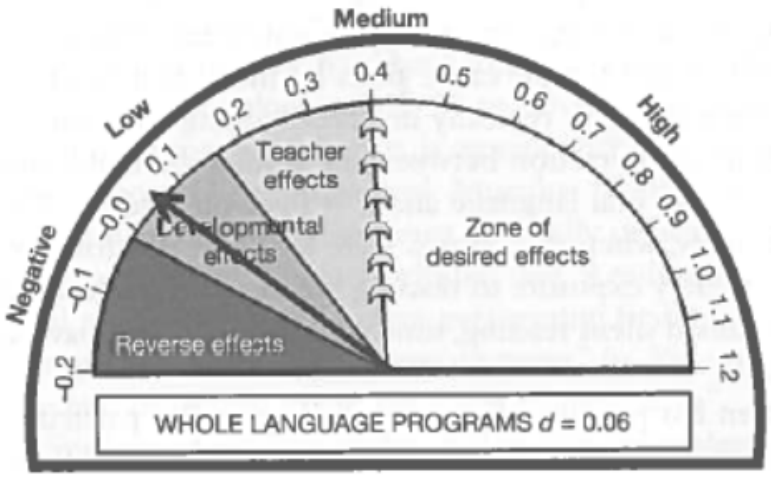


- Cargo-Cult science (Feynman): when something looks like an experiment, but it isn't.
    - If is is a quasi-experimental design, or action research, it's not actual experimental research. Quasi-experimental designs are good for what they do, but they have their place. They could be a first study, to see if it's worth it to do something more controllable.
    - If it's not random assignment and maintain baseline equivalence. it's problematic.
    - Most brain research is quasi experimental
        - example: when taking two groups of kids and have them engage in reading but one group has LD, and their brains are measures, that's not random assignment, that's quasi experimental.

00:23:00 Objective investigation and accumulation

- Meta-analyses are a great way to get an overview of the entire literature
    - there are some useful sources out there

00:23:36 Example - Whole Language Programs

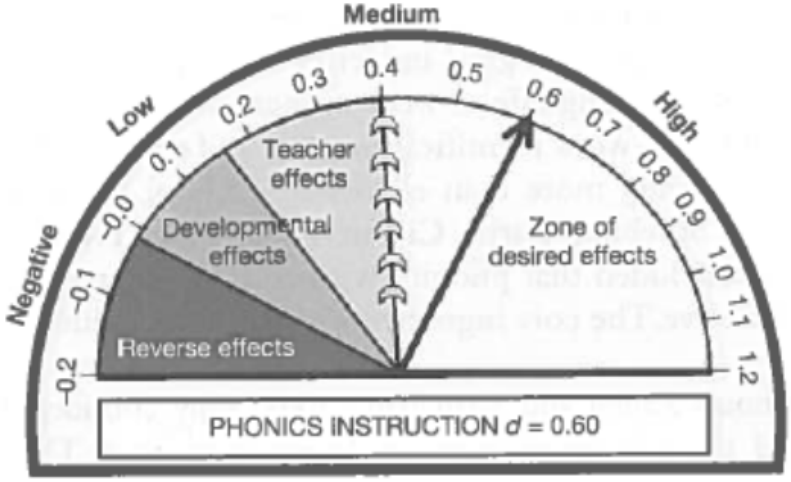


KEY	
Standard error	0.056 (Medium)
Rank	129th
Number of meta-analyses	4
Number of studies	64
Number of effects	197
Number of people (1)	630

We know that this doesn't work, because

- 4 meta analyses
- 64 studies
- 197 effect sizes
- 630 participants
- average effect size d = 0.06 (almost zero)

00:23:36 Example - Phonics instruction

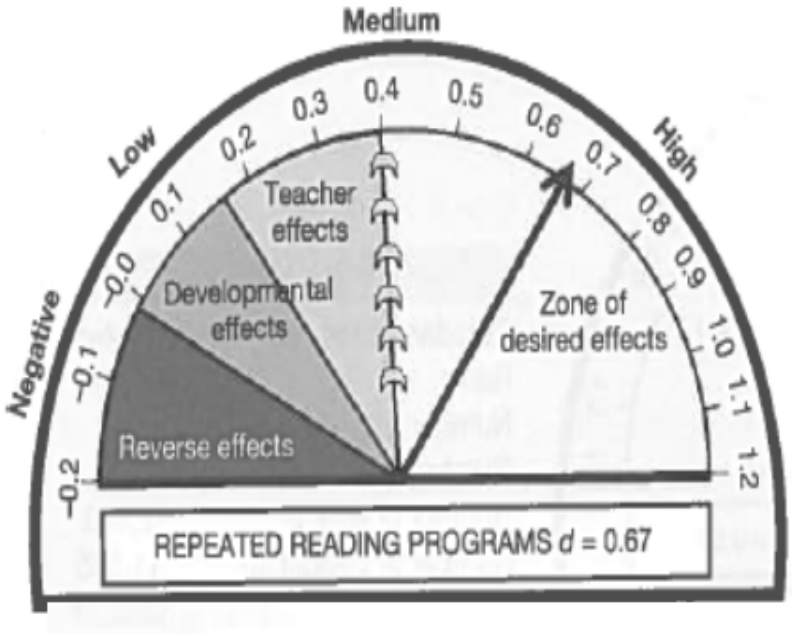


KEY	
Standard error	0.221 (High)
Rank	22nd
Number of meta-analyses	14
Number of studies	425
Number of effects	5,968
Number of people (5)	12,124

We know that it works, because

- 14 meta analyses
- 425 studies
- 5968 effect sizes
- 12124 participants
- average effect size d = 0.6 (medium to high)

00:24:25 Example - Repeated Reading

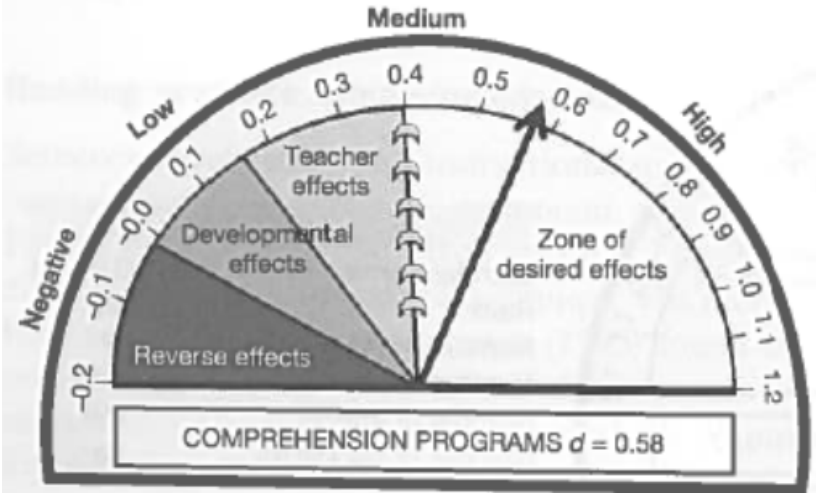


KEY	
Standard error	0.080 (High)
Rank	16th
Number of meta-analyses	2
Number of studies	54
Number of effects	156
Number of people (0)	na

We know that it works, because

- 2 meta analyses
- 54 studies
- 156 effect sizes
- n/a participants (not reported)
- average effect size d = 0.67 (medium to high)

00:24:40 Example - Comprehension Programs



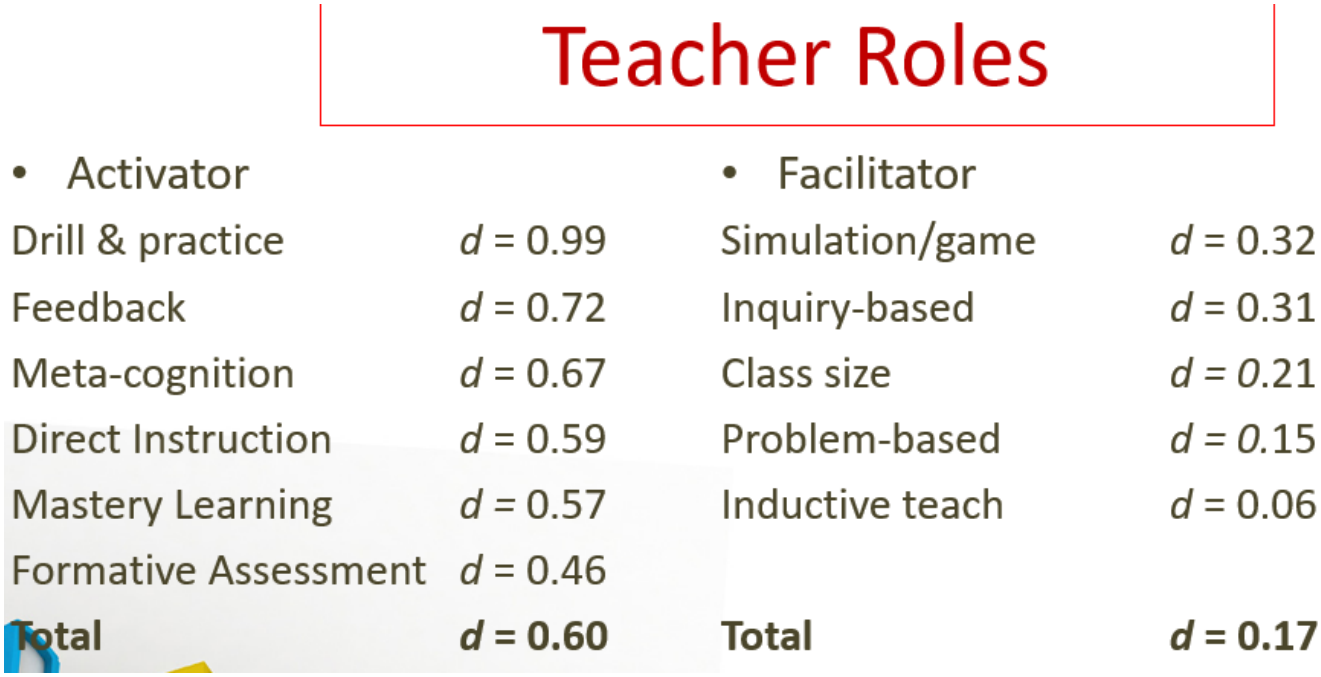
KEY	
Standard error	0.056 (Medium)
Rank	28th
Number of meta-analyses	9
Number of studies	415
Number of effects	2,653
Number of people (6)	11,585

We know that it works, because

- 9 meta analyses
- 215 studies
- 2653 effect sizes

- 11585 participants)
- average effect size d = 0.58 (medium to high)

00:25:20 Teacher roles



Teacher as Activator

As shown in the slide, teacher as activator, where the teacher is actively engaged in in teaching, we notice that

- **drill and practice** shows the largest effects. Practicing works. It helps synthesize and generalize things better.
- **feedback** - providing good instructional feedbackis one of the most important things we can provide.
- **Metacognition** high effects
- **Formative assessment** high effects
- **Direct instruction** high effects
  The overall effect for teacher as activator appears to be d = 0.6 which is a medium to large effect

Useful resources:

- Hattie (2009)

> - https://intensiveintervention.org/
> - https://ies.ed.gov/ncee/wwc/PracticeGuides

The intensive intervension website contains important resources that link to the original studies, provides a comparison chart that lists the adequacy of the study design, and the effect size. In addition, each study is linked to a page that describes the sudy in greater detail, including the price and the associated research.

What they have done well at ies is these practice guides - they get researchers together, they look at all the research, do a meta-analysis in order to determine what the is evidence based that works and here is how to implement them, tailored to practitioners, all free.

Examples incluse, one on

- teaching comprehension to younger children
- teaching math to girls
- turning around low performing schools

At the intervention website, you can filter, for example, by reading, and this should tell you the rating, and the effect sizes

---

So far we discussed:

- how to evaluate experimental designs
- how to collect this accumulation of studies

Now let's focus on the application of these resources.

There are many things that a re commonly done but lack a good research base.

Common Practices without Research

00:29:30
Example: Guided reading,
- this from page 96 of
https://www2.gov.bc.ca/assets/gov/education/early-learning/teach/resources/primary_prog.pdf

What does the research on guided reading show?

- There are no experimental studies on guided reading
- Iaquinta (2006) for example provides 22 references, 18 (82%) books/position, 1 qualitative dissertation, Juel (1998), NRP (2000), Torgeson (1998)

They concluded that there were no experimental studies done in guided reading.

In other words, it is not a closed and shut case, but rather that we lack the data in support of effectiveness, or more transparently, it lacks research base.

Example: differentiated instruction
https://journals.sagepub.com/doi/abs/10.3102/0034654320933536
Few studies showed *how* to differentiate instruction

00:32:24
Example: multiple intelligences

Howard Gardner's research
https://www.semanticscholar.org/paper/The-Effect-of-Multiple-Intelligences-Theory-Based-%3A-Bas/ded6bc3b8492044ad419661a338668ae693f7a92

- Bas (2016) 75 studies
- d = 1.08
- All conducted in Turkey
- 64 master's theses and 11 dissertations
- 0 examined reading

There are no published studies in this meta-analysis, and none looked at reading, so at the very learst it lacks research base for reading.

00:33:15
Example: Levelled Literacy Intervention
(Fountas & Pinnell)

- Most common reading program (Edweek Research Center, 2019) 43% in the US
- Ransford-Kaldon et al. (2010)
- University of Memphis
- Kindergarten d = 0.09
- First-Grade d = 0.32
- Second-Grade d = 0.16

00:35:36
Example: Orton Gillingham

- Ritchey & Goeke (2006) (meta analysis)
- 12 studies
- 5 reported OG > control/comparison
- 11 Quasi-experimental
- 1 experimental – college students, spelling with Wilson program (which is evidence based, but only with college kids and looking at spelling alone)
- 0 published studies since 2006
  In other words, it is not a closed and shut case, but rather that we lack the data in support of effectiveness, or more transparently, it lacks research base.
  - in OG there's an emphasis to decoding, it's not themulti-sensory approach
    - and for decoding there's many other interventions that are just as good that would work just as well and cost less, but also have a more solid research base
    - refer to the links provided (with the comparison of the interventions) to see what works

---

00:37:00

What works

- Systematic phonics instruction
  - is critical in learning how to read
  - there are at least 20 citations on this, too many to list
- Cooperative learning
  - dozens of studies showing it to be effective
- Writing and reading should be integrated
  - Steve Graham paper https://doi.org/10.1002/rrq.332
- Spelling = decoding
  - lots of studies in support of it, spelling is a good indicator of decoding
  - if a kid can't spell, it's quite possible they can't decode, especially if they're struggling with reading
- Mechanics of writing are important
  - penmanship, spwlling, letter formation all of it influences writing

---

00:40:32
"Meta-analysis of targeted small-group reading interventions" - Hall & Burns (2018)
https://www.sciencedirect.com/science/article/pii/S0022440517301231

- 27 studies, g = 0.54
- All with a control group, 15 met WWC standards (0.59) and 7 met with reservations
- Targeting (k = 13, g = 0.65)
- Comprehensive (k = 14, g = 0.36)

00:41:44
What we know:

- reading fluency is related to comprehension
- repeated reading, although not as popular, has a strong research base (recent emta0analysis finding large effects)
    - on the other hand, partner-reading, or reading widely, as commonly known, independent reading, the effects for that are d=0.18
    - so again repeated reading was better than some of the other approaches
    - readworks.org is a fascinating resource, if you have not encountered it yet. You can sort by type, by grade, thousands of passages available. And they have published some national norms based on several sources of data.

00:43:44 commentary on repeated reading falling out of favour

- it is true that the kid needs phonemic awareness (re OG)
- fluency without phonemic awareness is not going to help
- if a kid needs help decoding, phonemic awareness is porbably not going to help
- Burns: I work backwards based on the skill level
    - if a kid needs phonemic awareness, that's a priority
    - if a kid needs decoding (if phonemic awareness is fine), that's what they get
    - if a kid needs fluency, where their decoding and phonemic awareness are fine, they get fluency, and so on
    - **There is no such thing as universal intervention**
        - you have to target the student's deficit and address the intervention to match that
        - the website shown (slide 24) is the one with the most recent data on Oral Reading Fluency https://files.eric.ed.gov/fulltext/ED605146.pdf

---

00:44:53

Questions that remain: further research needed

- Catts & Kamhi (2017) team do some interesting research on whether we can teach reading and comprehension
- sidenote: an interesting CATTS paper, recent https://ila.onlinelibrary.wiley.com/doi/abs/10.1002/rrq.352
- the paper referred to: https://doi.org/10.1044/2017_LSHSS-16-0033

The question asked by researchers: san we actually teach reading comprehension. Is that really what we teach, or is it that we are teaching the component parts, but comprehension helps. In other words, is comprehension a byproduct of good reading.

Cromley & Azevedo (2007) did an excellent piece of research https://psycnet.apa.org/doiLanding?doi=10.1037%2F0022-0663.99.2.311 one of my favourite studies in 2007- Burns says - where they wanted to see what leads to comprehension. They identified five things, and the first two were bacground knowledge and vocabulary, by far the most important two things, number three was the correct inferences about reading, number four was word reading skill and number five was strategy use.

So if a kid is struggling with comprehension, we might do a strategy, typically. But according to this paper, it might be better to instead teach the kid how to read the words.

Also I would argue - Burns says - that if a kid is struggling with comprehension, he is also struggling with fluency.

Then two studies, one published in psychology in schools and the other in assessment for effective intervention,

where we found that - says Burns - that unless a kid can read at least 60wpm, comprehension does not occur. So most of the time when we see a comprehension deficit it is really a fluency problem.

So based on this information - Burns says - the best thing we can do is to reinforce the bacground knowledge and vocabulary. Yes, we should teach strategy, and yes we should work out word reading skills, yes we should teach inferencing, but if I want a kid who is struggling with comprehension to do better, the background knowledge and vocabulary are sometimes the way to go.

00:46:42 Content Areas - Promoting Acceleration of Comprehension and Content

- i.e. science, social studies
    - these areas should be flooded with science of reading
    - most of the time, at least 50% of the time, especially in high school, these interventions occur in a content area course
- PACT - Promoting Acceleration of Comprehension and Content Through Text
    - g = 0.59 for Social Studies (Swanson et al., 2017)
    - small effects on reading comprehension
    - so if the kid has poor decoding skills, this intervention probably won't help them, but if the decoding and the fluency skills are there, then this intervention is probably appropriate

---

00:48:00 Project-based learning
This is a wonderful study, very impressive.

- Project-based learning (Duke et al., 2020).
- Reading for an authentic purpose.
    - this may be over-simplified. It's more than kids just reading for authentic purpose. But remember SOR is not just about decoding. Most people who vie for authentic reading cringe at the concept of authentic purpose, yet there's several studies out there, beautiful randomized designs with large N, randomly assignmed by teacher, really well done. They found an average effect size of 0.4 for social studies, and an effect size of 0.1 for inferential reading.
- Social Studies = 0.48
  Informational Reading = 0.18

00:49:17 Phonemic Awareness effects with older kids is still something that is in need of further research.

Obviously, if a kid needs PA that's what they should get, whoever, I maintain - Burns says - that PA is not their core deficit.

> These are some quotes from the National Reading Panel
>
> - The effect of PA training was by far the largest for students in preschool, but that
> - "kindergarten was significantly larger than the effects in 1st grade and in 2nd through 6th grades."

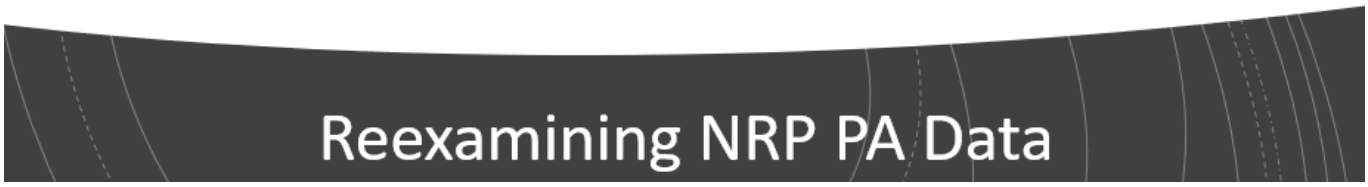
- "These findings indicate that younger students gained the most PA."
- "Transfer of PA to spelling was greater among kindergarteners than among 2nd graders. There was NO TRANSFER TO SPELLING (emphasis added) among the 2nd through 6th graders for whom the effect size did not differ from zero."

00:50:35

Means and Ranges of Effect Sizes by Reading Outcome Measure

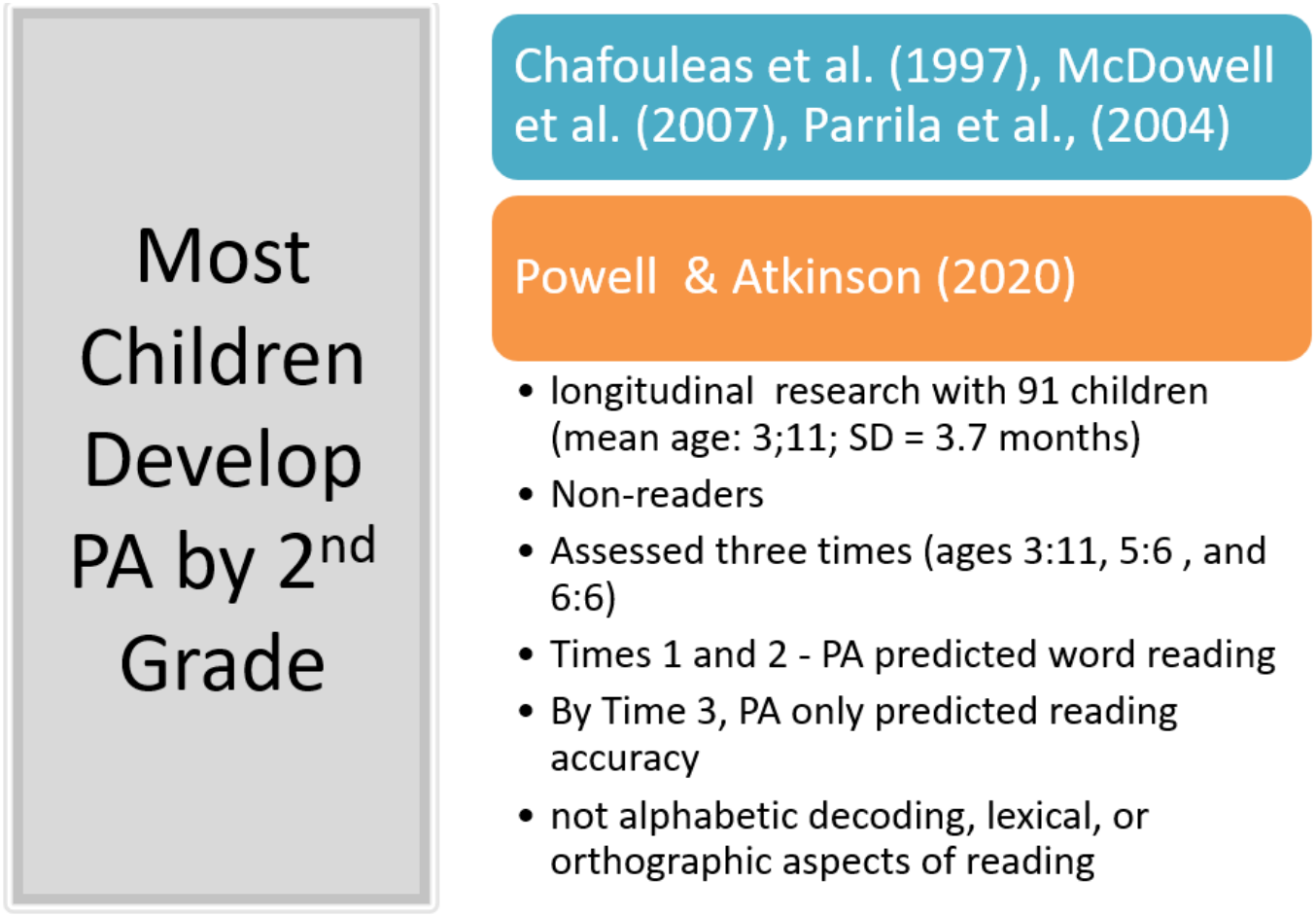
	N	Mean ES	*SD*	Minimum	Maximum
Pseudowords	24	.84	.80	−.19	3.60
Words in isolation	48	.92	.89	−.05	4.33
Contextual reading	24	.37	.38	−.37	1.18

(Burns, 2003)



Note how connected reading had a much smaller effect size, relatively speaking.

00:51:38

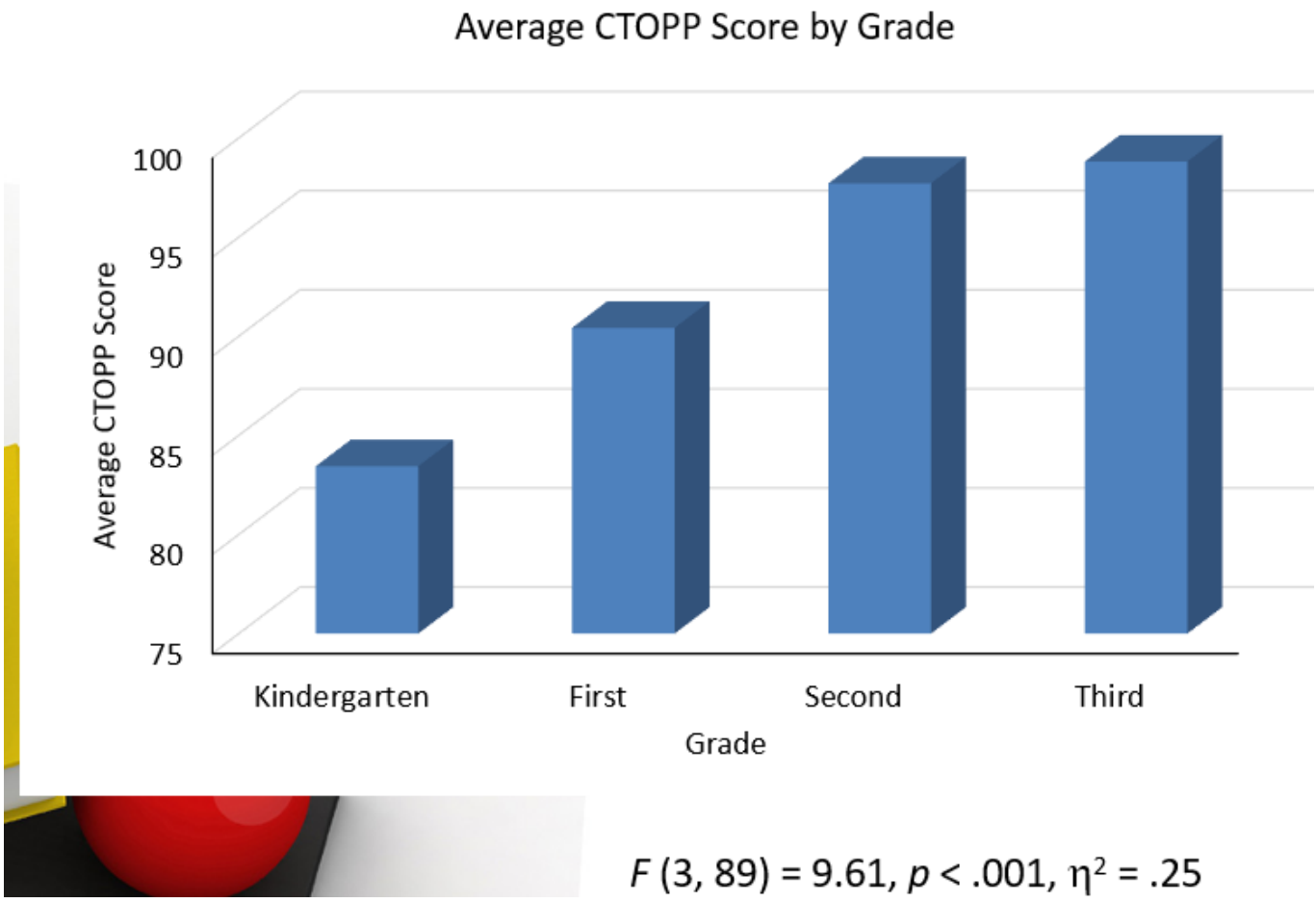


- The first square: they say that most children develop phonemic awareness by second grade.
- The second square, a longitudinal study with 91 children, where the children were about 4 years old on average (3y11m).
- All these were struggling readers, check out how by year 3 PA only predicted reading accuracy, and not alphabetic decoding, or orthographic aspects of reading

---

00:52:57

- Reanalysis of screening study
- 93 struggling readers



$F\ (3, 89) = 9.61, p < .001, \eta^2 = .25$

Upon reanalysis - completed by Burns himself, reanalysing the data - by looking at 93 children, just struggling readers, this is the average CTOP phonemic awareness composite score.

- Remember, this is an age-based standard score,so a score of 3 in kindergarten for example is entirely different from a score of 3 in 3rd grade.
- So this is the age-based standard score, so obviously
  - the kidergarteners were all below the 15th percentile on average,
  - the first graders were slightly below the 25th percentile
  - but by second and third grade their agerage score was aboe 95, getting close to the 95th percentail of average score.

00:53:50 So if you look at the relationship between them

Relationship Between DIBELS Composite and CTOPP Score			
Grade	*N*	Correlation	Number of Students Low PA
Kindergarten	18	.35*	15 (83%)
First Grade	19	.19	10 (53%)
Second Grade	24	.27	6 (25%)
Third Grade	32	.02	5 (16%)

This goes to show that by 3rd grade, phonemic awareness intervention will probably not help 27 out of 32 kids, not that PA is not important for those that need it.

00:55:00 Question about repeated reading, the long term effects.

There's a wonderful meta-analysis from 2016 where they found longitudinal effects (Suggate, 2016)

> Suggate, S. P. (2014). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77-96. https://doi.org/10.1177/0022219414528540

Here they found longitudinal effects of

- fluency intervention over time
- decoding intervention over time
- and through fluency intervention over time, comprehension as well
- result: all of them go down over time, but they can still see meaningful effects longitudinally
  Burns: we don't look at maintenance and generalization of interventions often enough. This is a legitimate concern

---

00:56:00

- PA is critical in kindergarden through 1st grade
- Haggerty - according to Burns - is a terrible intervention, because it tries to teach too much and doesn't really focus enough, but it's a wonderful supplemental tool for K-1 if you're not teaching phonemic awareness. However, there are no published studies as of yes that I know of reasonably well.
- Bottom line: **If you 're not teaching PA, you need to be teaching PA for grades K-1, but by 2nd grade, or by 3rd grade for sure, the number fo kids that this will be the right intervention for is going to go way down, and we don't have enough data for 3rd grade, unfortunately**

00:57:11 What about vocabulary?
The meta-analysis from the National Reading Panel only have four areas, not five. Vocabulary and comprehension usually appear together because it is tough to tease those two things apart.

00:57:46 From a SOR this is critical

> Ellis, A. K., & Bond, J. B. (2016). Research on educational innovations. https://doi.org/10.4324/9781315617145

Three levels of research (Ellis & Bond, 2016)

- Level I – Consistency with theory
    - Is it consistent with theory: there is a reason why good solid interventions with a good solid base of research don't maintain over time
    - 1. they're too hard to do and no one does it right
    - 2. they stray from the theory that they developed from
    - when this happens, it does not work anymore
    - consistency with theory does matter
- Level II – Does it work (tightly controlled)
    - this is the tightly controlled study
- Level III – Does it work in the real world
    - does it generalize, does it work in the real world
    - so, we need a lot more translational research
    - we know more about reading and less about how to teach reading (Shanahan, 2020)

00:59:20
Teachers are not taught SOR

- a survey revealed nearly 80% of teachers and educators define phonemic awareness as sound-symbol relationship

    > - https://trace.tennessee.edu/cgi/viewcontent.cgi?article=6743&context=utk_graddiss

- *Phonemic awareness is the ability to hear, identify and manipulate individual speech sounds iin words*

    > - https://www.youtube.com/watch?v=kRUXxERF3SY

01:07:35
SLP are a great resource for phonemic awareness related questions. In fact most of the research on PA stems from SLP research.

01:08:10 What level does PA need to be for a student to be able to effectively engage in the reading process?

- Burns: I argue you need to be good all the way to manipulating. You need to have that whole continuum. Unfortunately, for many kids that continuum is pretty sequential. So if a kid can't manipulate, they don't have PA. All the way to manipulation, the entire continuum has to be in place for us to be able to say that yes, the child has PA.

01:09:50 At what point do you know enough phonemic awareness rules?

- Burns: it's such a developmental sequence. If a child can't manipulate phonemes in 5th grade, that's a problem. For decoding, the answer is a bit more difficult.

01:16:10 Case of low processing speed and its impact on meeting a fluency goal, or let's not make fluency a goal because it's not fair to said student, given they have low

processing speed.

- Burns: processing speed is ione of the measures that translates the least to instruction, and it makes the most intuitive sense that this is the case. Now, what it does tell you is that it is an area of remediation: so if a kid is low on processing speed, you may want to focus more on being fluent in skill, although, the effects of doing so are pretty small. So as a school psychologist I would not write a recommendation that a kid do repeated reading if they have slow processing. But if a kid scores low in fluency that may be the case for repeated reading as a recommendation.
- I wouldn't put too much stock in a score as a prediction, ie. case study, this child (alias Lonny) who scored really low in K-1 and the school did not want to work with him anymore and we said we still want to, and the mom absolutely wanted to and by 2nd grade he was a grade level peer. In fact, a study we did with Sara Schoen (2012?) we saw that pre-intervention data such as IQ predict post intervention reading quite well, but they do a very poor job at predicting both. So you just can't tell ahead of time who will respond to intervention and who won't.

01:18:25 Matt Resources
https://skippmizzou.weebly.com/resources.html

01:20:40 on 'Stealth Dyslexia'
If a child is a good reader, but has spelling issues in 5th or 6th grade, it may not matter as much as when a child has speling issues in 2nd grade, because this can potentially mean that there's some issue that may not show up until later in 3rd or 4th grade.
So K-1 if everything is good except for spelling, then it is a red flag.

01:23:05 on late readers, and possible heredity (Dev Hx - Developmental History)
The single best predictor of reading problems of any kid's recordis a family history of reading problems. But in a study of over a hundred of LD reports we reviewed, none of them reported a family history of reading difficulties. But if parents had reading difficulties but are now professionals and doing fine, I am not going to use that as a reason not to intervene.

01:24:40 When kids are in High School can't decode and nobody recognized it
Burns: that should never happen. That's inexcusable. It does happen, but it shouldn't We know so much about reading, we screen so much, that should not be happening.

The curriculum, instruction and environment are just as important as the learner in meeting outcomes
We need more instruction as opposed to labels, especially when IEPs end at some point, but the identification does not mean the kid is going to get improved instruction and improved outcomes
Tier 2 interventions: addressing student needs are just as important, despite parents preferring that the schools teach reading

---