

# Research for TPR Storytelling

Reproduced by permission of Sky Oaks Productions, Inc., the publisher of *Learning Another Language Through Actions*, 6th Edition-Year 2000 by James J. Asher

## *Is there any research to support the effectiveness of TPR Storytelling?*

Yes, there is. Todd McKay developed new products called TPR Storytelling. McKay furnished me with data from his students and asked me for a statistical analysis to determine the effectiveness of the storytelling approach.

## *Student Groups*

A class of 30 middle school students who experienced TPR Storytelling (TPRS) were compared with a class of 30 students in a traditional Audio-Lingual Method (ALM) class. Both classes were exposed to the same set of vocabulary. Then both classes listened to a story none of the students had heard before, but the story contained familiar vocabulary.

## *Results*

On a ten item true-false test to assess student comprehension of the “novel” story (one they had never heard before) the TPRS students had significantly higher comprehension compared with the ALM students. The TPRS students had a mean of 7.6 and a standard deviation of 1.83 compared with the ALM students who had a mean of 5.83 and a standard deviation of 1.88. A  $t$  test for independent samples yielded a  $t$  of 3.69, which was significant at  $p < .001$  for 58 df. (Note:  $p < .001$  means that there is less than one chance in a thousand that we made a mistake in concluding that in the “population,” the average performance of the TPRS students will be higher than the average performance of the ALM students.)

## *Effect Size (ES)*

Jacob Cohen from New York University published through Academic Press, the book: *Statistical Power Analysis for the Behavioral Sciences*, 1969. According to Cohen, “effect size” is the variance in the dependent variable that is “explained” by the independent variable. Remember that a “significance” test merely tells us that one group on the average is *different* from a comparison group. Effect size ( $r^2$ ) gives us an indication as to the magnitude of the difference. For data collected on human subjects, a small ES is .02, a medium ES is .12 and a large ES is .25. For my students in statistics, I recommend that if the significance test is “significant,” always follow-up by finding the effect size and reporting this information to the reader.

## *Effect size in the McKay study*

In the McKay study, the independent variable was the instructional strategy of TPRS compared with ALM. The dependent variable was the true-false test for understanding a “novel” story. The effect size of  $r^2 = .19$  tells us that the independent variable of TPRS had a substantial impact on the dependent variable which was student performance on the ten item true-false test. Since the dependent variable had a low ceiling (of only ten items), it seems to me that the ES would be dramatically larger if the true-false test was administered for multiple stories instead of just one story.

# Research for TPR Storytelling

## *Recommendations for follow-up studies*

For graduate students who would like to expand upon this pilot study to create an exciting master's thesis or doctoral dissertation, here are some suggestions:

### *Your student groups*

Be sure that the students in each group are comparable in age, aptitude and hours of exposure to instruction in a language program.

### *Use Multiple Stories*

Use multiple stories rather than only one story so that the ceiling is high enough for differences in performance to show up between the groups. I recommend that the stories be "novel"—ones that the students have not heard before but contain familiar vocabulary that they have experienced in their classroom instruction.

McKay's books for Year 1, Year 2, and Year 3 have built in "novel" stories called Main Stories which would be ideal in future research studies. McKay prepares students for a "Main Story" with four short stories illustrated with cartoons that contain all the vocabulary the student will hear in the Main Story.

After students hear a Main Story for the first time, measure their listening comprehension with the ten item true-false questions which McKay provides. There are nine Main Stories. Plot a curve showing the performance of your students on each of the nine Main Stories. This is an impressive display to show parents and administrators.

### *Assessing listening comprehension*

Assess listening comprehension by playing either to the left brain or to the right brain. Here is how to do it: For the left brain, ask a set of true and false questions about each story. For the right brain, ask the students in the experimental and comparison groups to draw some pictures that illustrate what happened in each story. Code the drawings in some way so you know which group they came from.

Recruit two impartial judges who independently look at a drawing selected from the experimental group and another drawing selected from the comparison group.

Instruct the judges: "Is **Drawing A** compared with **Drawing B** better on "story understanding" or is **B** better than **A**?"

Next, have the two judges look at two more drawing without consulting each other. For example, they compare **Drawing A** and **Drawing C**. Then they look at **Drawing A** and **Drawing D**, and so forth. Deciding between only two items at a time is the simplest decision-making one can ask of a judge.

Scoring "story understanding" is simple: Which group (experimental or comparison) had the most drawings selected? To interpret the results, I recommend a statistical procedure called 2 x 2 chi square (in which the expected frequencies by chance are 50:50).

### *Assessing reading skill*

Students either listen to a Main Story for the first time or they read it. I suggest that every other story is for listening or reading. Either way, measure listening and reading using the left and right brain testing procedures I recommend above.

## **Research for TPR Storytelling**

### *Assessing writing skill*

Each student is given a printed set of familiar vocabulary items in the target language. They are asked to write an original story (the wilder and crazier the better) in a limited time period.

The assessment is a double-blind procedure. Two or more impartial language teachers shall look at each story. They shall not know which instructional group the student was in and they shall not know the identity of each student.

Ask the teachers independently to compare the stories two at a time ( i.e., A with B, A with C... etc. and make a simple decision such as which of the two is better on spelling? Then, compare the stories again for originality. Then compare again for grammar, and so forth.

### *Assessing speaking skill*

Give each student a list of familiar vocabulary in the target language and record on video each student telling a story that they made up in the target language using the list.

Again in a double-blind procedure ask two more impartial language teachers (who do not know the students) to compare two students at a time. Compare first on fluency. Then view the videos again and compare for originality, and so forth.

### *Reliability of the assessment measures*

After you score the stories for each student, determine the reliability of the teachers' judgements. This is usually a Pearson Product Moment Correlation for two judges (teachers) who independently evaluate each student in a group. This is critical because if reliability is unacceptable (i.e.,  $r = .69$  or less) then you should not take the next step which is to apply a "significance" test that will show which group excelled on a particular measurement. Every measurement must have acceptable reliability.

### *Significance tests*

As a rule of thumb, if your dependent variable (that is, your assessment) is continuous, then apply a  $t$  test if you only have two groups. The assessments I have suggested are continuous. If you matched students on age, aptitude and other variables, then use the  $t$  test for correlated samples. If you did not match, then use the  $t$  test for independent samples.

If you are comparing two or more groups, you may use analysis of variance if the samples are independent or analysis of covariance if the samples are correlated. Your left brain may be complaining that, "This looks complicated! I don't understand it! What is this all about?"

Your professors can advise you. If I can be of assistance with a specific question, please let me know. My e-mail is [tprworld@aol.com](mailto:tprworld@aol.com)

The Office of Education would be an ideal place to submit for a research grant to support this worthwhile project.

### *A note to graduate students from the editor,*

Dr. James J. Asher

Originator of the Total Physical Response, known worldwide as TPR

e-mail: [tprworld@aol.com](mailto:tprworld@aol.com)

web: [www.tpr-world.com](http://www.tpr-world.com)

## Research for TPR Storytelling

I receive many inquiries from graduate students who want to explore TPR in a master's thesis or a doctoral dissertation. The basic research showing the effectiveness of TPR has been thoroughly established years ago. I did this work in a series of research projects supported by grants awarded from the U.S. Office of Education, The U.S. Office of Naval Research, The Defense Department, the State of California, and San Jose State University. For a summary of this work, see my book, *Learning Another Language Through Actions*.

What remains to be explored are the parameters of TPR Storytelling (TPRS). We need carefully designed research studies to answer fundamental questions such as:

Is there a significant difference in storytelling performance between students who acquire vocabulary with classic TPR compared with students who acquire vocabulary with gestures only?

Is there a significant difference in performance between students who experience stories that are exaggerated, bizarre, and surprising compared with stories that are mundane?

Is there a significant difference in performance for stories that are non-goal-directed compared with stories that are goal-directed, such as:

*How to give directions to a taxi driver.*

*How to buy a ticket on the train.*

*How to find your way to the hotel, restaurant, police station, etc.*

Is there a significant difference in storytelling performance between students in elementary, high school and college?

Is there a significant difference in performance between students who experience mini-stories compared with a standard length story?

How many stories are optimal before adaptation sets in? (Adaptation may be measured by student resistance as indicated by remarks such as, "Please, not another story?" "Can't we do something else today?" etc.

What is the optimal mix between classical TPR, storytelling and other linguistic tools such as grammar explanations, patterned drills, etc.?

How do storytelling students perform on standardized proficiency tests? Do they outperform students in traditional classes? If so, by how much?

What are the correlations between predictors such as academic aptitude, school grades, age, socio-economic status, etc., and the criterion of performance as a result of storytelling?

*Note:* Performance can be measured in short-term retention, long-term retention, and attitude ratings by students. Performance can also be assessed by ratings of proficiency in speaking, reading, and writing by teachers who do not know what kind of training each student has experienced.

I can see scores of exciting research projects for a master's thesis or a doctoral dissertation focused on developing scientific answers to these important questions about TPR Storytelling.