# HYBRID TASKS: PROMOTING STATISTICAL THINKING AND CRITICAL THINKING THROUGH THE SAME MATHEMATICAL ACTIVITIES

Einav Aizikovitsh-Udi, David Clarke, and Sebastian Kuntze

*Even though statistical thinking and critical thinking appear to have strong links from a theoretical point of view, empirical research into the intersections and potential interrelatedness of these aspects of competence is scarce. Our research suggests that thinking skills in both areas may be interdependent. Given this interconnection, it should be possible to stimulate both forms of thinking through the one task. This paper explores the implications of an exploratory qualitative study into processes when working on tasks encompassing both these areas for the design of tasks that simultaneously stimulate critical thinking and domain-specific thinking.*

*Keywords*: Critical thinking; Hybrid tasks; Statistical thinking; Task design

Tareas híbridas: promover el pensamiento estadístico y el pensamiento crítico por medio de las mismas actividades matemáticas

*Aunque desde un punto de vista teórico, el pensamiento estadístico y el pensamiento crítico parecen tener fuertes conexiones, la investigación empírica sobre la intersección y la potencial interrelación entre estos aspectos de competencia es escasa. Nuestra investigación sugiere que las habilidades de pensamiento en ambas áreas pueden ser interdependientes. Dada esta interconexión, debería ser posible estimular ambas formas de pensamiento mediante una tarea. Este artículo explora las implicaciones de un estudio cualitativo exploratorio sobre los procesos de pensamiento en tareas que abordan ambas áreas para el diseño de tareas que simultáneamente estimulan el pensamiento crítico y pensamiento específico de un dominio.*

*Términos clave*: Diseño de tareas; Pensamiento crítico; Pensamiento estadístico; Tareas híbridas

Aizikovitsh-Udi, E., Clarke, D., & Kuntze, S. (2014). Hybrid tasks: Promoting statistical thinking and critical thinking through the sema mathematical activities. *PNA*, *8*(3), 115-126.

Mathematics classroom instruction is generally organized around and delivered through students' activities on mathematical tasks (Doyle, 1988). Notably, in all of the seven countries that participated in the TIMSS 1999 Video Study, eighth-grade mathematics was most commonly taught by spending at least 80% of lesson time in mathematics classrooms working on mathematical tasks (Gallimore, Garnier, Givvin, Hollingsworth, Jacobs, Chiu, et al., 2003). For our purposes, complete description of a mathematical task requires specification of the intentions, actions and interpretations of both teacher and student/s, together with details of the context in which the task was undertaken and by whom (Mesiti & Clarke, 2010). In this paper, task mostly refers to the written stimulus, although we also provide the details required to understand the performative realisation of the task as a mathematical activity. Classroom activities are coherent actions shaped by the instructional context, in general, and, in particular, by what is taught through the use of tasks (Stodolsky, 1988). The tasks that teachers assign can determine how students come to understand what is taught. In other words, tasks serve as a context for students' thinking, during and after instruction. Doyle (1983) argues the point that "tasks influence learners by directing their attention to particular aspects of content and by specifying ways of processing information" (p. 161).

To achieve quality mathematics instruction, the role of mathematical tasks to stimulate students' cognitive processes is crucial (Hiebert & Wearne, 1993). Contemporary curricula prioritise more than just procedural knowledge. In the mathematics curricula of countries as culturally and geographically distant as China, Australia and Finland, student thinking is promoted in more sophisticated terms than simply knowledge of facts and procedures. Korea and Singapore, in particular, seek to retain the high level of mathematical competence documented in TIMSS and PISA results, while improving higher order thinking and problem solving expertise. This poses the question as to whether both can be achieved. The possibility is explored in this paper that the same tasks might be employed to realise both goals. The tasks employed came from the research of Kuntze and his colleagues, the actual interviews were conducted by Aizikovitsh-Udi, the analyses of the task responses with respect to statistical thinking and critical thinking were coordinated by Kuntze and Aizikovitsh-Udi and the task-specific interpretive analysis was done by Clarke.

In summary, the centrality of tasks in mathematics classroom is evident from theoretical perspectives as well as in empirical results from international comparative studies. The role of mathematical tasks provides a key to any attempt to understand teaching and learning in research on classroom practices in mathematics. But can a given task stimulate and promote both discipline-specific thinking and more generic forms of higher-order thinking? To explore this question, we take Statistical Thinking (ST) and Critical Thinking (CT) as our exemplars of the two modes of thought.

## Statistical Thinking and Critical Thinking

In a well-known definition of statistical literacy by Gal (2004), a "critical stance" is included among the key attitudes for successful CT-hence, Gal includes such attitudes in his definition of statistical literacy (Reading, 2002; Wallman, 1993; Watson, 1997). However, being critical in statistical contexts is not only an attitude. It is possible to describe specific abilities that have to be used in order to critically evaluate statistical data. Two key concepts or overarching ideas in CT relevant for a critical evaluation of data are manipulation of data by reduction (Kröpfl, Peschek, & Schneider, 2000) and dealing with statistical variation (e.g., Watson & Callingham, 2003). Successfully manipulating data by reduction requires the awareness of such things as that calculating a mean value affords an overview on the original data, but it reduces the initial information. Hence, the resultant statistical value is (only) an indicator corresponding to a specific mathematical model, and we should not forget that it reflects only a part of the information. In order to critically evaluate the data, we might need additional information about the distribution, such as the variance, or information about extreme values.

CT skills rely on self-regulation of the thinking processes, construction of meaning, and detection of patterns in supposedly disorganized structures (Ennis, 1989). CT tends to be complex and often terminates in multiple solutions, each with advantages and disadvantages, rather than a single clear solution. It requires the use of multiple, sometimes mutually contradictory criteria, and frequently concludes with uncertainty. This description of CT already suggests links with ST, such as dealing with uncertainty, contradictions and a critical evaluation of given claims (McPeck, 1981). Dealing critically with information—a crucial aspect for both domains—demands critical/evaluative thinking based on rational thinking processes and decisions (Aizikovitsh-Udi, 2012; Aizikovitsh-Udi & Amit, 2008). Can a single task be used to elicit and promote both forms of thinking?

In-depth analyses are required into how CT and ST may interdepend. In order to design the type of hybrid tasks proposed here, it is essential that we understand the connection between the two forms of thinking that provide the specific goal for the use of such tasks in mathematics classrooms and how these connect to task characteristics.

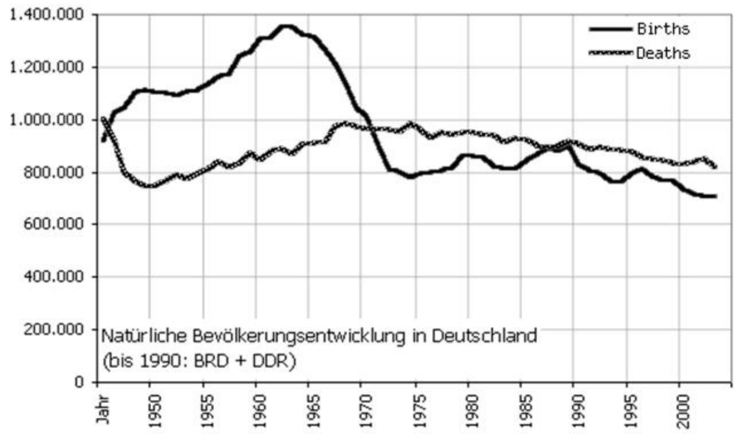
## Investigating Hybrid Tasks

In attempting to stimulate particular thinking skills, it might be thought best to target either discipline-specific thinking (e.g., twelve numbers have a mean of 10 and a standard deviation of 2, what might the numbers be?) or generic CT skills in discipline-free contexts (e.g., five people are isolated by flood, under what circumstances would it not be appropriate to share the available food equally?). Hybrid tasks seek to promote both. This paper explores the actions such tasks promote, the rationale for their use, and their design characteristics.

In order to explore thinking processes related to tasks in the domains of both ST and CT, individual semi-structured interviews were conducted with mathematics teachers. By using mathematics teachers as subjects, basic content competence can be assumed and it becomes possible to examine their content-related higher order thinking skills, both in terms of CT and CT. The interviews focused on thinking-aloud when solving tasks and each lasted about 40–50 minutes. Beyond solving the tasks, the interviewees were also free to give their personal views on the tasks.

In the following section, the results from one interview with a single teacher are used to exemplify the sort of data generated and the type of task likely to stimulate ST and CT. The analysis concentrated on identifying ST and CT as employed by the interviewees. Our interest is not just in the capacity of a single task to stimulate both ST and CT, but whether the interaction between ST and CT made them mutually supportive or disruptive. A first analysis was done focusing on ST only (Watson, 1997), then a second analysis employed a CT point of view (Ennis & Millman, 2005). We then carried out a combined interpretative analysis in order to examine relationships between CT and ST elements. In this methodological approach, the analyses were done by two coders working in parallel. Figure 1 shows the first task analysed.

„The Germans don't have enough children", was a recent newspaper's headline. Consider the following diagram about the development of the German population:




From which year on has the population been decreasing?

*Figure 1*. Task "population" (Kuntze, Lindmeier, & Reiss, 2008)

Nena was an experienced US secondary mathematics teacher with 20 years of mathematics teaching experience. She was dedicated to improving her teaching and had been participating in a year-long professional development program for mathematics teachers at the time this interview occurred. In the interview, Nena was asked to solve the problem in Figure 1, while thinking aloud.

*Interviewer:* What do you think? From which year on has the population been decreasing?

*Nena:* Well…the population has been decreasing since 1963 until about 1973, where the population begins to rise a bit...but it is still down since 1963...

*Interviewer:* Are you confident about it?

*Nena:* Of course… I am sure! You can look at this problem mathematically, anywhere there is a positive slope you could say the population is increasing, where there is a negative slope the population is decreasing. But compared to 1963, it's always been down.

Dealing with assumptions is one of the key elements of CT (Ennis & Millham, 2005). Nena initially assumed that the graph of births completely determined the population development. Even when asked to reflect on this assumption, Nena did not generate possible counter-arguments for testing her initial assumption, nor did she appear to question this assumption. She tended to seek confirming evidence rather than evidence that might challenge her initial assumption.

Seen from the perspective of ST, Nena chose an inappropriate statistical model for interpreting the data given in the diagram (description of this task in Kuntze et al., 2008). She appeared to focus on the data related to the births only, and she deduced her conclusions from a mathematical consideration of slopes. Even when encouraged by the interviewer, she did not check this model against the full data given in the diagram.

Nena's answers show deficits both in CT and ST. Looking at the relationships between CT and ST: At the very beginning, Nena shows only a partial perception of the evidence, focusing on the birth data from 1963 on. This selective focus may have been a result of the headline given in the diagram ("The Germans don't have enough children"). This headline may have triggered Nena's misinterpretation of the births as determining the population, from the turning point in 1963. It is interesting that Nena emphasised that it was possible to "look at this problem mathematically", which suggests that she saw a discrepancy between looking at the situation from the perspective of a mathematical model and looking at it from the perspective of the context (population and children born in Germany). Possibly the mathematical or statistical model is taken as an authority that is used to justify the appropriateness of the assumption instead of questioning the model chosen initially. It is possible to conclude from this combined analysis that the elements of reasoning in both domains interfere and interact. Given this interconnectedness, help in either domain, either in CT and

ST, may have had a positive impact on the thinking process as a whole. Moreover, the CT and ST perspectives offer not only simultaneous and parallel ways of interpreting the reasoning process, but, through a combined analysis, can explain how CT and ST can be mutually beneficial, reinforcing related reasoning approaches. Tasks that stimulate the use of both CT and ST are consequently of practical importance.

However, CT and ST are not always interdependent in an obvious way, as the following example suggests (Figure 2 and the corresponding interview section).

> Mrs. Blum would like to buy a reliable Laptop, either a C-Pad or an S-Top.
> In a computer magazine, 400 laptops of each brand have been tested. In this comparison the C-Pad has turned out to be more reliable.
> In the evening she talks to three friends. Two have S-Tops and never had problems. The third had a C-Pad, but had so many hardware problems with it, that he has sold it again immediately.
>
> With which of the following statements do you agree?
>
> □ Mrs. Blum should buy an S-Top, because the friend with the C-Pad had made bad experiences, whereas the friends are happy with the S-Tops.
> □ Mrs. Blum should buy a C-Pad, because the test in the computer magazine is based on a high number of computers, not only on one or two.
> □ No matter how she decides, it can happen that she gets a Laptop that causes problems frequently.

*Figure 2.* Task "Laptops"

*Interviewer:* So, what do you think, with which of the following statements do you agree?

*Nena:* I would agree with the third statement.

*Interviewer:* Can you explain, please?

*Nena:* Yes...400 computers is not a large sample when talking about computers so I would go with the third statement and just listening to the comments of her friends and the consumer's magazine, it is possible that both the computers could be just so.

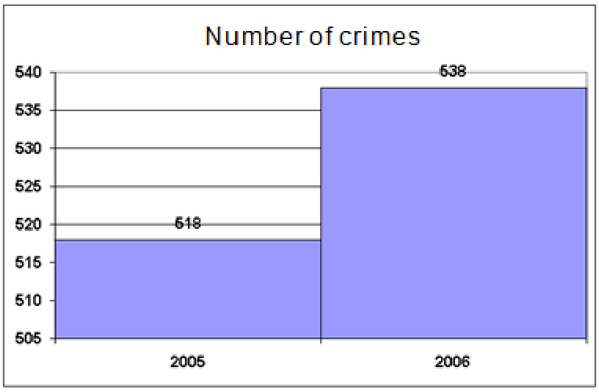
*Interviewer:* Are you sure?

*Nena:* Sure. I like this question!

Seen from the perspective of CT, Nena questions not only the experiences of the friends, but also the results from the study with the 400 laptops. On this basis, she expressed agreement only with the third statement, which highlights questioning evidence as a sub-aspect of CT. However, considering Nena's answers under the lens of ST, she appears not to acknowledge the statistical power of the sample of 400 laptops. Nena remarks that it is not possible to make a prediction on the base of the data, and she appears to compare the number of the 400 laptops to the number of all laptops. She does not reflect in depth on the statistical power of the magazine study.

From a joint perspective, Nena's dominant critical attitude may have blocked her use of elements of ST, for example reasoning related to the sample size and representativeness. This example gives insight into how CT and ST may interfere. In this case, CT practice acted to the detriment of ST. Conversely, ST can be dominant over CT, as the following example (associated to the problem in Figure 3) suggests.

The police president shows the following diagram and says: "This diagram shows, that since 2005, the number of crimes in the city center has increased, so that we have to expect that it will further go up in the next years. We need more policemen for patrol in the city center."




In a more extensive press communication, Fred has found the data of the last 7 years:

| Year | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 |
|---|---|---|---|---|---|---|---|
| Number of crimes | 504 | 528 | 525 | 499 | 529 | 518 | 538 |

Do you agree with the interpretation of the police president?

Why or why not?

*Figure 3.* Task "Crimes"

*Interviewer:* …and here...Do you agree with the interpretation of the police president? Why or why not?

*Nena:* I do not see a big difference with the number of crimes for any given year that would warrant extra police force to be hired...

*Interviewer:* Can you justify it, please?

*Nena:* It looks like the average is approximately 520, which is close to all the numbers so I do not think anything different is really happening from any given year.

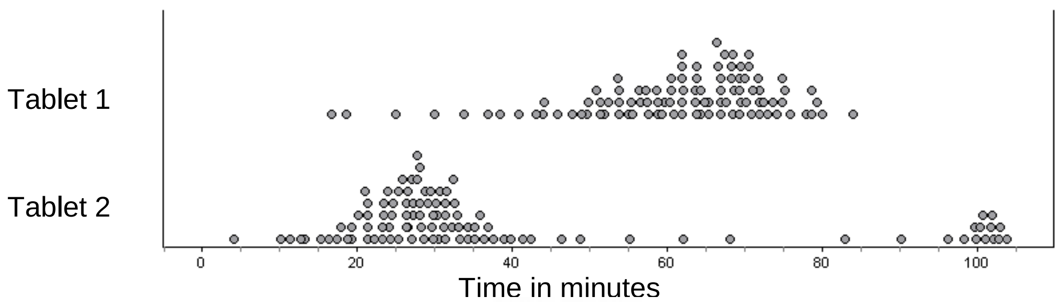
According to Ennis' taxonomy (1989), one crucial element of CT is raising questions, having doubts, and exploring key definitions, like "Crimes". In this task, the nature of the crimes is a key consideration. For example, if we knew that all the crimes were murders, we might decide differently than if the crimes related to paying taxes or fraud. No question about such a definition was raised by Nena and the focus was purely numerical. In this task, Nena employed a restricted set of CT skills.

From the point of view of ST, Nena used an appropriate model and showed an awareness of statistical variation. By these means, she arrived at the conclusion, that, given the variation of the data of the past years, the rise of the crime number is not significant. Consequently, seen from the ST perspective, Nena showed an appropriate understanding of the statistical situation.

Looking at this part of the interview in the joint CT and ST interpretation mode, the analysis suggests that Nena successfully questioned the statement of the police president by using the data given in the problem and a statistical argument. She appeared to remain in the statistical domain, giving more details related to the model she had chosen (distances to the average value). This focus on ST may have hindered her use of any of the CT skills, such as analysing and questioning the definition of "crime" as the key notion here, questioning the evidence (i.e. the way the data had been collected), etc.

Questioning data plays a role also in the following example related to the problem in Figure 4.

A company produces two sorts of headache tablets. Both sorts have been tested in a laboratory with respectively 100 persons suffering from headache. The diagram below shows, how long it took until the headache was over. Each point represents one test person.




Dr. Green: **Tablet 1 is the better one!**

Dr. Jenkins: **Tablet 2 is the better one!**

Find counter-arguments!

No, because ________________

Find counter-arguments!

No, because ________________

*Figure 4.* Task "Tablets" (Kuntze et al., 2008)

*Nena:* And here…I think tablet one takes too long to get rid of the headache. Tablet two seems to get rid of the headache a lot quicker for the majority of the people. You don't really know the age or weight of the people. Many factors play into the reason why a headache might occur so the statistics are poor. Based on the chart, I would have to pick tablet number two in hope of a speedy recovery.

From the perspective of CT, Nena not only evaluates the given statements, but she also shows CT elements when going beyond the data given: She gives examples of relevant influencing factors, and questions the data provided in the diagram ("the statistics are poor").

From the point of view of ST, the analysis of Nena's short answer yields that Nena chose an appropriate model and was aware of the key elements of the problem, even if she did not explicitly discuss the minority of cases with very slow recovery for tablet 2. These considerations led to her personal conclusion to pick tablet number two, as she obviously sees the chance of a "speedy recovery" as more important than the risk of a very slow recovery.

Looking at both CT and ST, the example appears to highlight how elements of CT can contribute to ST, for example when evaluating data, its presentation and analysis, planning data collection, etc. In the example, Nena suggests an analysis that takes into account the age or weight of the persons in the study. Conversely, aspects of ST like dealing with statistical variation and uncertainty can contribute to CT, especially when it comes to decisions in non-determinist situations, where full data is unavailable. These examples are intended to illustrate how both ST and CT skills can be evoked by the same task. We suggest that this models authentic and useful thinking practice more effectively than a more closed task that stimulated only CT and the application of taught procedures. Our question now, is "what are the messages for task design?"

## Conclusions: the Feasibility, Utility and Design of Hybrid Tasks

The tasks used in this exploratory study have certain distinctive characteristics and we would argue that each of these characteristics constitutes a key principle of task design:

- each uses a "real-world" situation as its "figurative context" (see Clarke & Helme, 1998);
- each provides succinct statistical information relevant to that context (Kuntze et al., 2008);
- the problem is stated very simply;
- some form of evaluation is integral to the problem (Ennis & Millham, 2005); and

- the task affords many reasoning approaches.

The exploratory study demonstrates that connections clearly exist between CT and CT at the level of individual reasoning practices. In seeking to stimulate both forms of thinking we suggest that an individual employing CT has access to a structured framework of analytical principles that guide and support their reasoning. That is, the relationship between measures of central tendency and variance, for example, structure any consideration of distribution of data that might be invoked in drawing evidence-based conclusions or making evidence-based judgements. On the other hand, the components of CT are not related in such a structured fashion and an individual's inclination to employ one strategy (e.g., questioning evidence or questioning assumptions) can be given expression without any obligation to also invoke other components of CT. Some CT skills resemble the heuristics that were the focus of the enthusiasm for problem solving in the 1980s and 1990s (Clarke, Goos, & Morony, 2007). Catalogues of such heuristics were similarly fragmented.

Ennis and others have catalogued CT skills (Ennis, 1989) and even arranged these categories in a form of hierarchy, but the connection between specific CT skills is under-theorised in comparison with CT. Nonetheless, the forms of CT identified in such classificatory schemes are clearly of significance, both as aspects of reasoning and as potential curriculum content. If it were possible to develop a structure for CT in which the component elements were not only identified, but also their relationship established, then to invoke one aspect of CT would serve to catalyse the use of other related aspects, because the connections between elements would be well known and understood. The question of how best to conceptualise these skills, how to integrate or connect them with other curricular goals, and how best to promote them and nurture their development in the classroom has been a major challenge. An earlier study (Aizikovitsh-Udi, 2012), using similar tasks, has documented efforts to produce CT through a program of instructional immersion in the related topic of probability. In this paper, we argue that particular tasks can stimulate the use, promotion and development of both Statistical and CT. We would like to suggest that an instructional program of hybrid tasks could be devised that provides the opportunity to employ CT, while simultaneously introducing students to the practices and structure of CT. The design characteristics of such hybrid tasks have been identified.

## References

Aizikovitsh-Udi, E. (2012). *Developing critical thinking through probability models, intuitive judgments and decision-making under uncertainty.* Published doctoral dissertation. Saarbrucken, Germany: LAP Lambert Academic Publishing.

Aizikovitsh-Udi, E., & Amit, M. (2008). Developing critical thinking in probability lesson. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano, & A. Sepúlveda (Eds.), *Proceedings of the Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 9-13). Morelia, México: Universidad Michoacana de San Nicolás de Hidalgo.

Clarke, D. J., Goos, M., & Morony, W. (2007). Problem solving and working mathematically: An Australian perspective. *Zentralblatt für Didaktik der Mathematik (ZDM International Journal of Mathematics Education), 39*, 475-490.

Clarke, D. J., & Helme, S. (1998). Context as construction. In O. Bjorkqvist (Ed.), *Mathematics teaching from a constructivist point of view* (pp. 129-147). Vasa, Finland: Faculty of Education, Abo Akademi University.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53*, 159-199.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*, 167-180.

Ennis, R. R. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18*, 4-10.

Ennis, R. H., & Millman, J. (2005). *Cornell critical thinking test, level Z* (5th ed.). Seaside, CA: The CT Company.

Gal, I. (2004). Statistical literacy, meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Dordrecht, The Netherlands: Kluwer.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: NCES.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*, 393-425.

Kröpfl, B., Peschek, W., & Schneider, E. (2000). Stochastik in der Schule: globale ideen, lokale bedeutungen, zentrale tätigkeiten. [Stochastics in school: Global ideas, local significances, central activities]. *Mathematica Didactica, 23*(2), 25-57.

Kuntze, S., Lindmeier, A., & Reiss, K. (2008). "Using models and representations in statistical contexts" as a sub-competency of statistical literacy–Results from three empirical studies. In R. Biehler & M. Shaughnessy (Eds.), *Proceedings of the 11th International Congress on Mathematics Education*. Monterrey, Mexico: International Congress on Mathematical Education.

Mesiti, C., & Clarke, D. J. (2010). A functional analysis of mathematical tasks in China, Japan, Sweden, Australia and the U.S.A: Voice and agency. In Y. Shimizu, B. Kaur, R. Huang, & D. J. Clarke (Eds.), *Mathematical tasks in* *classrooms around the world* (pp. 185-216). Rotterdam, The Netherlands: Sense Publishers.

McPeck, J. (1981). *Critical thinking and education.* New York, NY: St. Martin's Press.

Reading, C. (2002), Profile for statistical understanding. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town, South Africa: International Association for Statistical Education.

Stodolsky, S. (1988). *The subject matters: Classroom activity in math and social studies*. Chicago, IL: The University of Chicago Press.

Wallman, K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, *88*(421), 1-8.

Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). Israel/United States: IOS Press.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3-46.

This document was originally published as Aizikovitsh-Udi, E., Clarke, D., & Kuntze, S. (2013). Hybrid tasks: Promoting statistical thinking and critical thinking through the same mathematical activities. In C. Margolinas (Ed.), *PTask desing in mathematics education. Proceedings of ICMI Study 22* (Vol. 1, pp. 453-461). Oxford, United Kingdom: International Commission on Mathematics Instruction.

Einav Aizikovitsh-Udi
Beit Berl College
eia994@mail.harvard.edu

David Clarke
University of Melbourne
d.clarke@unimelb.edu.au

Sebastian Kuntze
Ludwigsburg University of Education
kuntze@ph-ludwigsburg.de

Received: December 2013. Accepted: January 2014.
Handle: