Science and Engineering Ethics (2019) 25:1261–1269
https://doi.org/10.1007/s11948-019-00088-1

OPINION/EDITORIAL

Purple Dragons and Yellow Toadstools a Versatile Exercise for Introducing Students to Negotiated Consensus

Brian P. Coppola[1] · **India C. Plough**[2] · **Huai Sun**[3]

Received: 13 December 2018 / Accepted: 28 January 2019 / Published online: 6 February 2019

Abstract

An activity called *Purple Dragons and Yellow Toadstools*, originally reported in 1987 as a training activity for jurors, was adapted as a priming exercise for a unit on teaching research ethics with undergraduate students. In this activity, learners develop skills for building negotiated consensus. The procedure involves individuals' ranking 10–15 moral transgressions and/or legal violations followed by a small group discussion in order to arrive at an agreed-upon ranking by the team. The framework has proved to be quite flexible, adaptable to different subject areas and with different populations of students.

Keywords Research ethics · Ethics training · Curriculum design · Teaching · Education

Introduction

Negotiated consensus is a hallmark pedagogical strategy in teaching research ethics (Cheruvalath 2017; Schrag 2005; Hess and Fore 2018; Bebeau et al. 1995; Davis 2011; Elliott and Stern 1997; Heitman 2002; Macrina 2005; Shamoo and Resnik 2014; Wueste 2005), particularly when using case studies, where students may need

✉ Brian P. Coppola
bcoppola@umich.edu

India C. Plough
ploughi@msu.edu

Huai Sun
huaisun@sjtu.edu.cn

[1] Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

[2] Residential College in the Arts and Humanities, Michigan State University, East Lansing, MI 48825, USA

[3] Department of Chemistry, Shanghai Jiao Tong University, Minhang District, Shanghai 200240, China

to reach agreement about an action when no clear-cut solution exists (Coppola and Smith 1996).

Negotiated consensus and ethical decision-making are significant educational outcomes in their own right, and fit well into the model of *Transformational Teaching* created by Slavich and Zimbardo (2012). In their model, Slavich and Zimbardo identify three "basic principles" that contribute to being an educated person: (1) conceptual mastery of the subject matter, (2) improved learning in how to learn, and (3) accruing learning-related attitudes, values, and beliefs. As psychologists, they define the last category as gains in constructs such as motivation and self-efficacy. This third category is also an entrée into a wealth of crosscutting outcomes in education, such as building a defensible argument, ethical reasoning, use of evidence, skepticism, pattern recognition, clear and coherent communication, and so on (Coppola 2013). These topics may not appear explicitly on a syllabus, yet they are implicitly critical features of disciplinary expertise and are called *disciplinary dispositions* (Coppola and Krajcik 2014).

Arguably, learning how to reach consensus in the face of individual disagreement is an important outcome in every discipline, in addition to being an explicit goal in teaching research ethics: balancing multiple perspectives, whether from evidence or from considering differing points of view, and considering effects of decisions on stakeholders. Throughout their education, learners are urged to arrive at a most defensible position, based on explicitly stated moral or theoretical frameworks, rather than to make the generally unlikely discovery of "a right answer" (Coppola and Smith 1996; Kovac and Coppola 2000).

We share here a warm-up exercise that we developed for use in a unit on teaching research ethics in chemistry courses (Coppola 2000; Coppola and Smith 1996), in both the United States and in China, as well as testing their use in an intercultural communication course in the humanities, where a comparable need exists to have students practice with negotiation in the face of potentially strongly held areas of disagreement.

Purple Dragons and Yellow Toadstools

While it is axiomatic that "people of good conscience can disagree," it is also true that there are times when people who disagree need to come to a consensus. One of us (BPC) encountered *Purple Dragons and Yellow Toadstools*[1] in 2005 while serving on a university level grievance board (the last stop for misconduct cases before the civil courts). This activity had been used for an unknown number of years as the training exercise for newly assembled members of the grievance board convened in

[1] Unfortunately, the origins of Purple Dragons and Yellow Toadstools, beyond "I think we used it in a workshop", are currently lost to time. The only citation to it is from Cordner and Brooks (1987) with a footnote that references an adaptation to the activity used by Smith (1974). In the 1987 report, the activity is used as a training exercise for jurors. While preparing this manuscript, both Cordner and Brooks were contacted, and neither could recall or find any more detail than what was reported in their article (the reference to Smith is a photocopy of a handout).

Table 1 Purple dragons and yellow toadstools (1987 and 2005 versions)

Purple dragons and yellow toadstools, 1987 (original version: Cordner and Brooks 1987, p. 37)	*Purple dragons and yellow toadstools*, 2005 The University of Michigan grievance board made slight changes to the 1987 version, as shown below:
The original checklist and set of instructions from 1987 follows here:	*Instructions:* Please read the following lists of potential violations and rank them in order from 1 to 15 (where 1 is the violation you feel is most severe and 15 is the least severe). There are no correct answers. Please be aware that there is not enough detail in these statements to make a truly informed decision. You will have the opportunity to compare and discuss your answers with others in your group and generate a group ranking.
Instructions: Rank the following violations of college rules according to the importance or seriousness you attach to them. Place a 1 in front of the most important, a 2 before the second most important, and so on. You have 10 min for this task. After members of your group have finished working individually, arrive at a rank order as a group. The group has 30 min for this task. Do not choose a formal leader, and try to arrive at your group ranking by consensus rather than by majority vote or by averaging scores.	A. Stealing a final exam and selling it to 15 students enrolled in that class
A. Cheating on an examination	B. Tampering with a colleagues experimental data
B. Defacing college property	C. Selling 2 grams of cocaine
C. Forging or altering an ID card	D. Lifting you mom's credit card and buying $200.00 worth of books and supplies
D. Distributing or selling drugs	E. Using the telephone to harass others
E. Plagiarizing on an English theme	F. Raping another University student
F. Punching a roommate during an argument	G. Plagiarizing sections of your term paper
G. Pulling a fire alarm as a prank	H. Providing alcohol to a minor
H. Smoking in a classroom with a posted "No Smoking" sign	I. Stealing a computer from the University
I. Making a false entry on an application	J. Physically threatening another student with a knife
J. Possessing marijuana	K. Removing a person, by force, and taking him or her to a location unknown to them.
K. Stealing a bicycle	L. Sexually harassing another student
L. Throwing firecrackers out a residence hall window	M. Dishonest reporting of investigative results
	N. Fabricating a letter of recommendation for admittance into graduate school
	O. Sending unauthorized e-mail communication to other students in the name of a faculty member

the graduate school at the University of Michigan (U–M). Table 1 shows two versions of the activity: the only published version, from 1987, and the version encountered by BPC as a training exercise in 2005.

The instructional design is as compelling as it is simple, and in our experience, it works as well for three people as for groups of at least 25–30. Individuals first rank a list of transgressions according to their severity; they then come together as a group (or subgroups) and enter into a conversation with team members to achieve a consensus ranking of the same items, where agreeing to disagree is not an option. The experience of ranking these items is viscerally effective: by using only a list of transgressions, every decision becomes a choice between the better of two evils. We immediately integrated the activity into our supplemental instruction program for first-year organic chemistry students (Varma-Nelson and Coppola 2005) as a priming exercise for a 6-week unit on research ethics (Coppola 2000). Subsequently, it was incorporated into a course on intercultural communication for arts and humanities students in the United States and also with groups of science students in China,

where there is an emerging interest in teaching research ethics in undergraduate and graduate education (Liao et al. 2018).

Implementation

At the University of Michigan, the context for teaching research ethics in which we have used *Purple Dragons* consistently, is with students who are enrolled in the Honors organic chemistry program, 2005 to date.[2] To compare its use in different contexts, we have also used it in an Honors chemistry program at Shanghai Jiao Tong University, and in an arts and humanities course at Michigan State University.

The broad outline of the *Purple Dragons* exercise is described to students prior to handing out the detailed information. Each student gets a handout with the same list of items. Without discussion, each person has 5–7 min to provide an absolute ranking (no ties) from 1 (most severe) to 15 (least severe) of the items, and to consider the reasons they are using to create their ranking. The rankings should be written down. Then, groups of 3–4 individuals, formed randomly, need to come to consensus on a ranking, which can take 20–45 min.

Purple Dragons and Yellow Toadstools, 2007

The 2005 version of the exercise was adapted so that items were directly relevant to undergraduate university students (Table 2).

After students have completed their individual rankings, the facilitator can face the board, away from the students (to preserve a degree of anonymity), and ask a volunteer to call out the number of hands raised in response to the question "who had Item A ranked as #1, #2, #3… etc." Individual responses can be aggregated on a 15×15 grid (e.g., 3 of 23 students ranked Item A as #5, and so on), prior to the group work, so that students can see the result from the individual deliberation.

A great variation is seen in the rankings, which we found to be typical, and provides grist for the subsequent consensus discussion, which has ranged from quiet and deliberative to "knock-down, drag-out" with varying levels of cooperation and intransigence. Even after discussion within the groups, variation in the rankings between the groups still exists, and this diversity of opinions was used as the context for a follow-up discussion that involved the entire class. The two prompts for that classwide reflective conversation are (a) what were the criteria and method(s) used by your group to come to its consensus, and (b) how and why did you arrive at them?

[2] A useful warm-up activity we developed for *Purple Dragons and Yellow Toadstools* is called "Legal versus Moral." A 2×2 grid is presented to students (x-axis labels: "legal" and "illegal;" y-axis labels: "moral" and "immoral") with the instruction to think about whether all four cells are occupied. 3–5 min of individual contemplation is followed by a period of the instructor gathering participants' ideas and prompting, as needed (e.g., where would you place humans killing other humans?").

Table 2 Purple dragons and yellow toadstools (2007 15-item version)

Purple dragons and yellow toadstools, 2007 Modified version of the 2005 items, adapted for undergraduate university students:
A. Stealing a final exam and selling it to 15 students enrolled in that class
B. Tampering with the video footage of a student's project on a computer in the SLC
C. Selling 2 grams of cocaine
D. Sexually assaulting a student on campus
E. Stealing a U–M ID with Blue Bucks and buying $200.00 worth of books and supplies
F. Using the telephone to harass others
G. Plagiarizing sections of a senior thesis
H. Providing alcohol to a minor
I. Stealing a computer from the University
J. Physically threatening a student with a knife
K. Removing a person, by force, and taking him or her to a location unknown to them
L. Sexually harassing another student
M. Dishonest reporting of investigative results
N. Fabricating a letter of recommendation for admittance into graduate school
O. Sending unauthorized e-mail communication to other students in the name of a faculty

Table 3 Purple Dragons and Yellow Toadstools (2015 10-item version)

Purple dragons and yellow toadstools, 2015 Abbreviated version of the 2007 items:
A: Steal an exam and sell it to other students
B: Dishonest reporting of experimental results
C: Sexually harassing another student
D: Fabricating your own letter of recommendation
E: Stealing a computer from the university
F: Selling 2 grams of cocaine
G: Plagiarizing sections of your term report
H: Raping another university student[a]
I: Stealing a credit card to buy books and supplies
J: Threatening another student with a knife

[a] The 2005 version, used in China, was administered prior to a 2015 revision, which changed "rape" to "sexual assault" because of possible trauma triggers. The version that US students received starting in 2015 replaced "rape" with "sexual assault" to comply with university requests for faculty to be cognizant of triggering statements.

Also in 2015, a 10-item version of the exercise was created and implemented in parallel to the 15-item version to determine if a shorter version (which saves time) would elicit comparable discussion and perceived value for the exercise (Table 3).

A different sub-set of first-year U–M organic chemistry students from the 2015 cohort completed the 10-Item exercise. A group of chemistry students in the Zhi-Yuan Honors College at Shanghai Jiao Tong University (SJTU), who were part of

a workshop being conducted by BPC, were also administered the 10-item version, again to make a baseline comparison with another population of students.

Individual and Group Deliberations

After the group deliberations, the class is brought together and students are asked what criteria they used individually to rank the transgressions. The time taken for this portion of the exercise ranges from 5 to 10 min. Typically, more time (15–25 min) is spent discussing group deliberations, as students often contrast and compare ideas that arise in the group discussion that are not used, or do not occur, to the individual.

Criteria for Ranking by Groups

In general, the criteria used by groups has been similar to those used in the individual rankings. In groups, the most common report for ranking was if the infraction was clearly breaking the law, and how serious that infraction was. Additionally, when breaking the law, actions putting one person (yourself) at risk were less heinous than when others and larger groups were put at risk; the permanence of the offense was also considered more serious than if the offense was reversible. In all of these populations, students mentioned the dilemma of needing to balance (a) how many people are affected, (b) whether the effects have lasting damage, and (c) the severity of physical or emotional/mental harm.

Group Strategies

Arguably the most important aspects of the *Purple Dragons* exercise are the strategies that the students report using, as well as the arising issues, when they need to come to consensus on their rankings.

In every setting, the discussion reveals the assumptions made by individuals about the interpretations of the offenses, and makes it clear that knowing more about the context matters. Students all note that an agreement about the context needs to be established for the sake of discussion, and that the contextual assumptions should be voiced and attached to the final ranking. Resolving the individual differences about the interpretation of context has probably as much value to the process of learning about negotiated agreement as the ranking itself (Itten 2017).

In these reported settings, group members went out of their way to engage all members of the group as a part of reaching a consensus, even when reaching that consensus was anywhere between difficult to contentious. Group dynamics emerged that reflect any group situation: animated, vocal (sometimes loud) debates and disagreements, contemplative and quiet conversations, a recurring use of phrases such as “now that is a good argument” or “I had not thought about it that way,” and even declarations of “I am changing my mind now.”

Because agreeing to disagree is not an option in this activity, the students invent strategies to try and move the debate forward, including (a) creating a set of binary

decisions ("A versus H") instead of holistic ones, (b) starting at opposite ends of the spectrum and working inwards, (c) creating ranked meta-groups based on general infraction (stealing, harming others) and then sub-sorting within the smaller groups, (d) going with a majority vote, particularly in the middle ranks, and (e) re-ranking the items as individuals (post discussion) and then using the average of the values.

Context may matter. This exercise was administered in different settings to see if there would be differences.[3] The way that US and Chinese students rank the item about selling cocaine stands out. In the follow up discussion, the US students generally contend that it is not actually harming anyone personally without them choosing to, saying that if one chooses to buy cocaine it is their decision. On the other hand, most hard drug offenses in China are enforced with severe punishment, including execution in the more egregious cases (Bi 2012).

Summary Discussion

Our impression as instructors follows. These exercises (*Purple Dragons*, and the moral/legal grid) have been used at U–M every fall term for almost 15 years, and they never fail to produce lively and engaged discussions among hundreds of students. Anedoctal feedback, as well as end of term surveys, indicate that students and instructors see the same compelling face value for why these exercises are interesting.

In our experience, both the 10-item and 15-item versions appear to accomplish the same results. The 15-item version is most likely better than the 10-item version at promoting a more extended discussion, whether deliberate or contentious, simply because there are more items in the "center" to consider. The 10-item version was used in both 2016 and 2017 at U–M, and the results were comparable to the 2015 implementation. We suspect that the structure of having groups of 3–4 items that are more clearly at each end of the spectrum is an unexpectedly positive feature of the exercise, because it gives the groups an early big-picture agreement before launching into the debate about those middle items.

The discussions that students have are on-task and relevant (legality versus morality; harm to self versus others; reversibility) and reflect common frameworks used in more formal discussions of moral philosophy (Rachels 1999). These concrete discussions might give theoretical discussions of topics such as utilitarianism or relativism a more meaningful context for subsequent instruction. Students are engaged; none is a bystander, waiting for the exercise to end.

The exercises described here are attractive because they require essentially no preparation, they have a high benefit/cost ratio, they are equally accessible across areas of study (and countries), and the items are easily customized to the local context. There is no absolute or one correct answer, after all, and so the adaptability is high. Educators who are developing instructional materials that promote more inclusive classrooms and other organizational structures (Gurin and Maxwell 2017)

[3] Preliminary data from 392 students in the three different settings is available upon request.

might rapidly integrate *Purple Dragons* as a way for students (and other participants) to take on the question of civil and negotiated discourse as a topic in its own right prior to other interventions and activities.

Acknowledgements The authors thank their students for continually engaging one another with enthusiasm and thoughtful discussion. BPC thanks the undergraduate group leaders who help implement these activities in their supplemental instruction sessions (L. Chen, L. Daboul, T. Friedlander, J. Gatti, J. Lawniczak, D. Luan, J. Luo, K. McKernan, A. Milen, A. Min, C. Nino, A. Nishii, P. Parker, M. Payne, M. Ryan, R. Tarnopol, A. Young). We dedicate this paper to the legacy of the unknown individual(s) who originally created *Purple Dragons and Yellow Toadstools*.

References

Bebeau, M. J., Pimple, K. D., Mustavitch, K. M. T., Borden, S. L., & Smith, D. H. (1995). *Moral reasoning in scientific research: Cases for teaching and assessment*. Bloomington, IN: Poynter Center for the Study of Ethics and American Institutions.

Bi, Y. (2012). On the death penalty for drug-related crime in China. *Human Rights and Drugs, 2*(1), 29–44.

Cheruvalath, R. (2017). Does studying 'ethics' improve engineering students' meta-moral cognitive skills? https://doi.org/10.1007/s11948-017-0009-x.

Coppola, B. P. (2000). Targeting entry points for ethics in chemistry teaching and learning. *Journal of Chemical Education, 77*(11), 1506–1511.

Coppola, B. P. (2013). The distinctiveness of a higher education. *Journal Chemical Education, 90*(8), 955–956.

Coppola, B. P., & Krajcik, J. S. (2014). Discipline-centered postsecondary education research: Distinctive targets, challenges and opportunities. *Journal of Research in Science Teaching, 51*(6), 679–693.

Coppola, B. P., & Smith, D. H. (1996). A case for ethics. *Journal of Chemical Education, 73*(1), 33–34.

Cordner, P., & Brooks, T. F. (1987). Training techniques for judicial systems. In W. Caruso & W. Travelstead (Eds.), *Enhancing campus judicial systems: New directions for student services* (pp. 31–42). San Francisco, CA: Jossey-Bass.

Davis, M. (2011). The usefulness of moral theory in teaching practical ethics: A reply to Gert and Harris. *Teaching Ethics, 12*(1), 51–60.

Elliott, D., & Stern, J. E. (Eds.). (1997). *Research ethics: A reader*. Hanover, NH: University Press of New England for the Institute for the Study of Applied and Professional Ethics at Dartmouth College.

Gurin, P., & Maxwell, K. (2017). Overview: Faculty development for inclusive educational environments. *Liberal Education, 103*(3/4), 6–9.

Heitman, E. (2002). Using cases in the study of ethics. In R. E. Bulger, E. Heitman, & S. J. Reiser (Eds.), *The ethical dimensions of the biological sciences* (2nd ed., pp. 349–352). Cambridge, MA: Cambridge University Press.

Hess, J. L., & Fore, G. (2018). A systematic literature review of us engineering ethics interventions. *Science and Engineering Ethics, 24,* 551–583.

Itten, A. (2017). Context and content toward consensus in public mediation. *Negotiation Journal, 33*(3), 185–211.

Kovac, J., & Coppola, B. P. (2000). Universities as moral communities. *Soundings: An Interdisciplinary Journal, 83,* 765–779.

Liao, Q.-J., Zhang, Y.-Y., Fan, Y. C., Zheng, M.-H., Bai, Y., Eslick, G. D., et al. (2018). Perceptions of chinese biomedical researchers towards academic misconduct: A comparison between 2015 and 2010. *Science and Engineering Ethics, 24,* 629–645.

Macrina, F. L. (2005). *Scientific integrity: Text and cases in responsible conduct of research* (3rd ed.). Washington, DC: ASM Press.

Rachels, J. (1999). *The elements of moral philosophy* (3rd ed.). Boston, MA: McGraw-Hill.

Schrag, B. (2005). Pedagogical objectives in teaching research ethics in science and engineering. *Science and Engineering Ethics, 11*(3), 347–366.

Shamoo, A. E., & Resnik, D. B. (2014). *Responsible conduct of research* (3rd ed.). Oxford, NY: Oxford University Press.
Slavich, G. M., & Zimbardo, P. G. (2012). Transformational teaching: Theoretical underpinnings, basic principles, and core methods. *Educational Psychology Review, 24*(4), 569–608.
Smith, A. (1974). *Purple dragons and yellow toadstools: A ranking experience*. Columbus, OH: Ohio State University.
Varma-Nelson, P., & Coppola, B. P. (2005). Team Learning. In N. Pienta, M. M. Cooper, & T. Greenbowe (Eds.), *Chemist's guide to effective teaching* (pp. 155–169). Saddle River, NJ: Pearson.
Wueste, D. E. (2005). A philosophical yet user-friendly framework for ethical decision making in critical care nursing. *Dimsensions of Critical Care Nursing, 24*(2), 70–79.