Writing and choosing problems for a popular high school mathematics competition

Robert Geretschlager · Lukas Donner

Accepted: 9 March 2022 / Published online: 29 March 2022
© The Author(s) 2022

Abstract
In this paper, we consider the issues involved in creating appropriate problems for a popular mathematics competition, and how such problems differ from problems typically encountered in a classroom. We discuss the differences and similarities in school curricula versus the generally agreed upon topics encountered in international competitions. The question of inspiration for the development of competition problems is dealt with from the standpoint of the problem author, while aspects related to the motivation of the contest participant, objective and subjective problem difficulty and mathematical precision in mathematics competitions are also discussed.

Keywords Mathematics competitions · Mathematical Kangaroo · Multiple choice problems · Popular mathematics

1 Introduction

The problems typically posed in popular or inclusive mathematics competitions are similar to, but not identical to, problems regularly considered in a typical mathematics classroom. The argument can be made that competition-style problems should be used in the classroom setting a bit more than is currently the case (see, for instance, Geretschläger, 2017, 2020; Applebaum & Gofman, 2015), but there are certainly marked differences. These result quite naturally from their differing goals. While many researchers are dealing with the goals of textbook problems from several perspectives, the situation is not as yet that well discussed for competition-style problems. Kontorovich (2015) asked several adults from the Competition Movement1 and reported that this group formulated four goals that can be achieved by competition problems. Besides learning meaningful mathematics, further pedagogical goals are the strengthening of students’ positive attitudes towards mathematics, the creation of cognitive challenges, and adding an element of surprise for students. Unlike a typical textbook problem, the intent behind a competition problem is neither to teach a foundational concept nor to check whether the concept has been understood or to solidify such comprehension. While these aspects may also apply to its consideration, the main reason for a competition problem to exist is motivational. When we are posing a problem for a competition with a large number of non-specialist participants, the most important aspect is always consideration of the question. Is this interesting? Will the problem, as set here, capture the imaginations of the students once they are exposed to it?

Of course, there is an underlying assumption that students willingly engage with all problems that come their way in school, but this is unfortunately not always the case. There are things that need to be understood and practised on a basic level before it is possible to move on to the applications that most people would consider ‘interesting’. On the other hand, it should be possible to assume that each problem posed in a competition is worthy of being thought about in its own right, simply because its solution is rewarding in its satisfaction of intellectual curiosity. A good competition question will always find a large number of participants invested enough to want to engage in finding its solution, simply because the problem is intrinsically of interest to them. Reznick (1994) called this phenomenon inevitability: once you see the problem, you feel you must solve it.

1 The adults of the ‘Competition Movement’ consist of competition designers and organisers and experts in problem posing.
The aim of this paper is to address the question of how to best create, refine and select such problems for inclusion in competitions. Currently, problem selection for competitions is not done with didactic research results in mind, and we strive to bring this element into play by emphasizing connections to current topics in education research, specifically with regard to word problem solving and problem posing. Furthermore, we summarize the state of the art in competition mathematics in order to offer the option to both the Competition Movement and the mathematics education community to build on this state for future research.

A specific challenge arises in selecting problems for a popular competition, where participants cannot be expected to be as self-motivated as for top-level competitions such as the Mathematical Olympiads (MO). The problems should be such that any reasonably receptive participant will want to engage with them, in the hope of experiencing the feeling resulting from a successful solution and therefore enriching their learning (Kenderov, 2006).

In the following, the matter of creating suitable problems will be considered through the lens of the International Mathematical Kangaroo (MK). One of us (Geretschläger) has been actively involved in the problem-selection process since 1998, and as group chair of the Student level (grades 11 and 12) for 20 of those years. Specifically, the discussions delineating the differences between potential problems to be chosen for this competition, from ‘textbook’ problems on the one extreme and ‘Olympiad’ problems on the other, have led to much fruitful reflection on this topic over the years. The research process for this paper was as follows: in a first phase, both authors, as actors and proven experts in the field, reflected on what they considered to be the most essential aspects in the creation and selection of tasks for the MK and agreed on a common list of factors, including curriculum, range of difficulty, precision, and packaging. In a second step, these aspects were reflected on in detail and linked to the current didactic discourse. The results of this two-step process are presented in Sects. 3 to 8.

The MK is an international competition that has been organised since 1991. Recently, it has attracted more than six million participants annually. It is organised in more than 70 countries, with this number growing consistently.

The competition is organised in the following six levels: Pre-Écolier for grades 1 and 2, Écolier for grades 3 and 4, Benjamin for grades 5 and 6, Cadet for grades 7 and 8, Junior for grades 9 and 10, and Student for grades 11 and 12 (and 13 where applicable). The number of items in each paper varies by level from 15 to 30, and the time available to solve them also varies by level from 60 to 75 min. Each level has an equal number of multiple-choice items (1 out of 5) worth 3, 4 or 5 points, with the relative difficulty of the items rising with the number of points available. One quarter of the available points are deducted for incorrect answers, with no penalties imposed for omissions.

The problems are created and selected by an international group of experts at an annual meeting, at which all participating countries are represented. They are specifically designed to provide an interesting challenge for all levels of knowledge and talent.

This paper is organized as follows.

In Sect. 2, we explicitly identify key differences between a competition problem and a textbook-problem. In this context, we compare problems of the two types with similar mathematical content. We also address the fact that the mathematical knowledge required to solve an item from the MK is often encountered in class at a younger age than that of the participants in the competition. In Sect. 3, the focus is on competition curriculum. The typical restrictions in the competition curriculum, as well as the possibly missing connection to mathematics taught in the classroom, are often cited as negative elements by members of the mathematics education community (Kenderov, 2006), and we address the resulting tension. Also, we discuss the general prohibition of electronic tools in competitions. This goes against the current vogue in general education in many countries, whereby practical calculation in regular mathematics classwork is usually farmed out to a calculating device, resulting in yet another nuance of difference in the mathematical content of competitions and the classroom standard. The main focus of Sects. 4–6 is the creation of appropriate, interesting and not too difficult high-quality problems and distractors for the MK. Simultaneously, we embed the practitioner’s considerations in the theory of problem posing. We discuss the question of where the inspiration for an interesting popular competition problem comes from. Also, we discuss the difficulty of finding good, easy and interesting problems for competitions, and deal with the fact that it is not as difficult to find ‘hard’ problems that are also interesting. The qualities required for a problem to be considered interesting are precisely what tend to make them hard. This is true for solutions requiring multiple steps, a combination of ideas from different topic areas, or open-ended options for solutions. In considering problems at this end of the spectrum, there are therefore other matters that come into play. This leads to the matter of when a ‘hard’ problem is too hard for a popular competition, and when it is still appropriate. More specifically, we consider the line separating appropriate problems for the MK from those of advanced competitions that require special preparation, like the MO. In addition to the consideration of the maximum and minimum levels of difficulty and complexity of an appropriate MK problem, there is also the matter of the structure of the problems. When creating problems for mathematical competitions, the problem poser must be aware of the fact that “such a problem should be well-formulated and have no unnecessary givens; its formulation
is supposed to be short and include some elements of innovation for a particular category of solvers” (Sharygin, 1991, cited by Kontorovich & Koichu, 2016). In Sects. 7 and 8 we address these aspects—specifically the aspects of the appropriate brevity of problem formulations (i.e., precision) and the elements of innovation of a problem (i.e., abstractness or ‘packaging’). We also establish connections to research findings on word-problem solving. This aspect reminds us that the MK contains a very special subset of these problems. Some problem posers argue that the packaging of a problem is of great importance for improving the attractiveness of a task in order to engage as many students as possible, and we address the relevance of such arguments in Sect. 8.

2 The essential difference between a textbook problem and a closely related competition problem

The following problems illustrate the difference between a good competition problem and an exercise in the classroom. Item A is taken from the Student paper of the MK 2018, where it was posed as a problem of medium difficulty. Item B is a problem for grade 9, taken from the textbook by Geretsläger et al. (2004). Both are problems from the same topic area, namely, triangles in the coordinate plane, and both deal with the mid-points of the sides of a triangle.

(A) The vertices of a triangle have the coordinates A(p, q), B(r, s) and C(t, u) as shown. The mid-points of the sides of the triangle are the points M(−2, 1), N(2, −1) and P(3, 2).

Determine the value of the expression \( p + q + r + s + t + u \).

(A) 2 (B) 52 (C) 3 (D) 5 (E) another value

(B) We are given a triangle ABC with vertices A(1, −3), B(3, 4), C(−1, 2). Calculate the common point of the three medians of the triangle.

Depending on the level of sophistication of a student solving the textbook problem B, the expected solution will be attained by calculating the averages of the x- and y-coordinates of the three points or as the common point of two medians. Either way, there is standard knowledge that students are expected to apply as studied in the classroom and explained by a teacher or in a textbook.

For the competition problem, the mid-points of triangle sides also play a role, but in an unexpected way. It cannot be assumed that students have been exposed to such a situation before, and yet it is not so hard for them to find connections to facts learned in the classroom, and apply them in an original way. If they notice that the coordinates of any of the given points M, N and P can be expressed as the means of the unknown coordinates of the triangle vertices A, B and C, they obtain

\[
p + q + r + s + t + u = \frac{p + q}{2} + \frac{r + s}{2} + \frac{t + u}{2} = -2 + 1 + 2 + (-1) + 3 + 2 = 5,
\]

which yields D as the correct answer. Nothing here could be considered mathematically difficult or sophisticated in any way. Nevertheless, the fact that the structure of the problem is of an unexpected nature makes the question a puzzle. Solving the puzzle requires an original idea. In this case, this is simply the collection of sums of the various unknowns in such a non-standard way that they can be replaced by known values. A student who finds this method, or some other original way to tackle the problem, has discovered something new for him- or herself.

This comparison gives an example of what is perhaps the defining difference between a competition problem and a classroom exercise. Mathematical tasks in classrooms aim at supporting students to engage in a range of mathematical activities with specific didactical goals, such as exploration, concept formation, or practising skills (Barzel et al., 2013). Many tasks are meant to be solved by applying learned facts in a standard way in a context of a type that has been previously encountered in order to practice. The competition problem, on the other hand, while also using the same basic tools, is meant to be solved by coming up with some (small) amount of original thought.

Certainly item A could conceivably be used in a classroom, but item B would typically not appear in any competition, as it does not offer a student outfitted with the requisite basic knowledge the opportunity to discover anything original.

Item A was presented for grades 11 and 12, whereas its content is typically taught in grades 9 or 10. A strong argument in favor of keeping the problems ‘age appropriate’ lies in the idea of making use of the tension between the recreational aspect of an interesting problem and its educational aspects. If students engage with a problem because they find it intrinsically interesting, there is a better chance that the learning process will yield the desired results. This could be interpreted to imply that it would be preferable to stick to
subject matter close to the curricular material students are studying at any given time in proposing competition problems, but this severely limits the options for competitions. It can reasonably be expected that students will find something interesting to engage with if the solution of a problem requires material that they are well familiar with, and not just material they are currently studying. It is a constant struggle to find an appropriate balance between these two viewpoints in choosing competition tasks.

3 The matter of curriculum

Topics chosen for competitions tend to be taken from limited areas of mathematics. Also, the type of reasoning and calculation expected in solving such problems tends to be similarly limited. Sharich (2017) pointed out that these phenomena imply a crisis of substance, i.e., “the lack of challenging new mathematics in problems offered at Olympiads” (p. 72). However, there is a wide international consensus for these limitations that has developed over the last few decades, resulting in something frequently referred to as “modern elementary mathematics”. It can be defined as the union of certain elementary methods and the problems that can be solved by means of these methods (see Koichu & Andžāns, 2009). Kenderov (2006) emphasised that especially through the Competition Movement, this important part of our mathematical heritage is kept alive and further developed. It is interesting and informative to take a look at the reasons why modern elementary mathematics is so widely accepted in the competition community.

First, there is the matter of diverging national school curricula. Reduction to the intersection of the various national curricula eliminates such topics as calculus, complex numbers, set theory or statistical analysis for international competitions, to name just a few. Unfortunately, finding this intersection is not as obvious as might be suspected, and students with a background in schools that teach a curriculum close to the agreed upon “competition curriculum” are at a competitive advantage in the world of the MO. Andreeescu et al. (2008) pointed out that what most of the countries that are successful at international competitions “have in common are rigorous national mathematics curricula along with cultures and educational systems that value, encourage, and support students who excel in mathematics” (p. 1251).

Next, there is the matter of technology. In contrast to the decades-long trend of including digital technology in curricula all over the world (e.g., Sinclair et al., 2009), there is a wide consensus in the Competition Movement that technology should not be permitted in most competition settings. The main reason for this is the fact that number structure and number properties (prime numbers, perfect squares or higher powers of integers, unusual types of fractions, etc.) are considered to be a fertile ground from which interesting logical puzzles can be harvested. Shi (2012) described how a classical problem on certain digits of higher powers of integers can be solved easily with the use of digital technology. Technology has the potential to open up new routes for students to construct and comprehend mathematical knowledge and new approaches to problem-solving (Bray & Tangney, 2017). However, by permitting the use of technology at competitions, the focus of attention would be shifted from mathematical reasoning towards the “mathematical digital competency” defined by Geraniou and Jankvist (2019).

There is a general feeling that this characteristic type of subtle thinking and reasoning in number theory is eliminated by hitting every number-based problem with the sledgehammer of technology.

As a third point, emphasis is given in competitions to thinking as opposed to calculation or direct application of learned facts. Clear preference is given to problems whose solutions require original insight on the part of the participant, while problems that rely heavily on facts and methods that can be studied in advance are pushed into the background. Methods relevant to the solution of problems in modern elementary mathematics are either not specific to any particular branch of mathematics (e.g., symmetry or equivalence), involve the analysis of very specific singular cases, or are very general methods from specific branches, such as the invariant method in combinatorics (see Engel, 1998 for a comprehensive collection of methods). One consequence of this feature is the fact that problems posed at competitions tend to be closer to puzzles in their structure than to applications to real-world situations.

The International Mathematical Olympiad (IMO) is often used as a benchmark for comparisons between competitions. This makes a lot of sense, since the IMO has the longest tradition of the high-level international competitions (since 1959), the largest number of participating countries (over 110 in recent years) and a wide level of acceptance. The problems developed for this specific competition are traditionally grouped into four categories, namely, Algebra, Combinatorics, Geometry and Number Theory (e.g. Djukić et al., 2011). While their meanings are subtly different from the standard usage in research mathematics, the consensus concerning these topic groups is quite clear.

Variations of these four IMO categories are often used to sort the problems posed in other competitions. At the MK, for instance, the categories used in problem selection are Algebra, Logic, Geometry and Number. While some of the most interesting problems do not slot readily into any of these categories, an attempt is usually made to have all four groups represented as equally as possible. This is, of course, due to the format of the competition, where many multiple-choice questions are to be answered under restrictive time constraints. The flavor of the questions should therefore be
as wide-ranging as possible, while also being accessible to as many participants as possible (Akveld et al., 2020).

This criterion means that problem-creation is restricted a bit. Of course, problems that do not require any knowledge at all, and can be solved with original clever ideas, are particularly esteemed.

4 Creating a competition problem

Empirical research on problem posing is comprised of three main perspectives, namely, focusing on problem posing as a cognitive activity, as a learning goal or as an instructional approach (Liljedahl & Cai, 2021). Although the term problem posing is used inconsistently in mathematics education research when focusing on situations within mathematics lessons (Baumanns & Rott, 2021), the term is less ambiguous when addressing the creation of competition problems. How one develops the capacity to pose good problems is the central question in research on problem posing as a goal of mathematics instruction (Cai & Leikin, 2020). Within a survey study, Lee (2020) found that mathematical experts are rarely involved in studies on problem posing. As Kontorovich and Koichu (2016) pointed out, there currently exist only a handful of papers devoted to the principles of experts’ problem posing for mathematics competitions. Research in this area seems to have started with two Russian papers in the 1990s by Sharygin (1991) and Konstantinov (1997). According to Sharygin (1991), there are six general techniques used for creating new problems, which he names Reformulating, Chaining, Considering special cases, Generalizing, Varying the givens and Inquiry. As Kontorovich and Koichu (2016) pointed out, however, being aware of these techniques and carrying them out is not enough for posing a problem of high quality, as other studies with students and mathematics teachers as problem posers show. One reason may be the pool of familiar problems that are an important component of an expert’s knowledge base (Sharygin, 1991), a fact that can be implicitly found in problem posers’ papers, such as those by Klamkin (1994) or Reznick (1994). An individual new to the process of problem posing may well be fully cognizant of the useful methods but may not have the breadth of familiar problems at their disposal to refer to. Elgrably and Leikin (2021) underpinned this assumption via use of Problem-Posing-through-Investigation (PPI) tasks, a mathematical activity that combines problem posing and problem solving. The authors compared the performance in PPI (in a dynamic geometry environment) of members (or candidates) of the Israeli IMO team and mathematics majors who excelled in university mathematics. One finding was that the former “were more fluent, flexible and original and produced more complex problems with more complex auxiliary constructions” although the latter “took

2 The author L. Donner was formerly known as L. Andritsch.
according to Smith (1982) or applying particular test-wiseness strategies as defined by Millman et al. (1965). These strategies give advantages in solving multiple-choice problems without actually applying thematic knowledge. By trying to offer particularly ‘beautiful’ and ‘suitable’ incorrect answer options, the problem poser gives students the option of bypassing the solving process in obtaining the correct answer, and although the design of the competition—lots of multiple-choice problems and not a lot of time to reflect on them—encourages students to guess when the odds seem advantageous, most problem posers would agree that this choice should always be based on content related ideas. The actual use of test-wiseness strategies by successful participants at the MK was studied by asking Austrian winners of the 2018 competition (Donner et al., 2021).

5 What makes interesting easy problems so hard to find?

It is quite a multi-faceted challenge to create a problem that is simultaneously interesting and easy. This is the case for competitions in general, but it is especially true for the MK.

One reason for this difficulty may be intrinsic to one of the points raised by Kontorovich and Koichu (2016). If the level of innovation of a created problem is too small, the poser is not surprised by the content of the problem and thus does not classify it as an interesting problem. There may therefore be a lack of motivation on the part of a potential poser to create an easy problem from such a starting point. Furthermore, the poser can only rely on a limited part of his pool of familiar problems for this purpose. The poser may nevertheless feel obligated to create an easy problem ‘here and now’ due to time constraints, or due to being asked to do so by his or her peers.

What exactly is meant by an ‘easy’ problem in the MK?

There is wide-spread agreement that the 3-point problems should be easily solvable by all participants in the MK (Akveld et al., 2020). These problems should be accessible to students not generally interested in mathematics. The whole idea of the contest is popularisation, and all students should therefore have a real chance of achieving a feeling of success after solving and understanding something.

Regarding complexity, there is a broad consensus in the MK, that 3-point problems should not require more than one idea to solve and should ideally require only a single step of calculation or logical reasoning. These aspects are illustrated by the solution of the following 3-point item from the Student paper 2020.

(C) The sum of five three-digit numbers is 2664, as shown on the board. What is the value of $A + B + C + D + E$?

This task can be solved by the fundamental idea of reformulation. A three-digit number $ABC$ can be written as $100 \cdot A + 10 \cdot B + C$. Applying the observation that each of the five digits occurs as a units- and hundreds-digit exactly once, the given equation can be reformulated and the solution found by one step of calculation: $111 \cdot (A + B + C + D + E) = 2664$ or $A + B + C + D + E = 24$.

An alternative is to note that the sum of all five digits ends on the digit 4, as seen in the unit-digit of the sum. This means that there is a carry-over of 2 in order for the tens-digit to result from the same sum, and the sum of the digits must therefore be 24.

Without the application of one of these fundamental ideas, a student will not be able to solve the problem easily. (Of course, another way to solve this problem is pure guessing of a possible assignment of the variables, but this approach will take much more time in general.)

Solving this prototype of an easy problem does indeed involve one single idea and one step of calculation. Nevertheless, it is quite far from a typical textbook problem.

The problem can be described as interesting, because although it is not possible to determine any of the indeterminants $A$ to $E$, their sum can be uniquely determined. (This basic idea occurs quite often in different problems spread out over all age categories and competition years.)

A second consensus broadly agreed upon within the MK community concerns the design of the 3-point problems. It is often the case in problem development for multiple-choice competitions that problems are designed in such a way that certain errors in reasoning or illegitimate simplifications will lead to one of the incorrect answer options. Such an offered solution is a trap for anyone following a specific incorrect (but anticipated) train of thought. This may very well be the intent of the problem poser, meant to make a problem interesting that seems to be a routine task at first glance. Such tactics are intentionally avoided in the 3-point section of the MK. Similarly, problems that involve distractors which ‘widen the spectrum’ of possible answers, like distractors of

\[ A B C + B C D + C D E + D E A + E A B = 2 6 6 4 \]

(A) 4 (B) 14 (C) 24 (D) 34 (E) 44

\[ (A + B + C + D + E) = 24 \]
Writing and choosing problems for a popular high school mathematics competition

the form ‘none of these’ or ‘more than $x$’, to questions about the number of possibilities for something, are avoided there. In general, such distractors can enrich the value of a problem and make it more interesting, because methods like ‘trying out’ distractors or finding $x$ possibilities for something cannot ensure finding the correct solution to the problem. This reduces the number of tactical options a potential solver has at their disposal in dealing with a problem. It follows that these kinds of distractors are not suitable for problems meant to be easy.

Finally, as students are generally unfamiliar with problems involving aspects of more than one content category, straddling the borders between categories automatically increases the level of difficulty. Therefore, in order to be ‘easy’, a problem should not involve more than one subject. Quite often, however, it is exactly such interaction between content categories that makes a problem interesting. This means that problems of this kind—such as example A in Sect. 2, which involves aspects of geometry and algebra—may not be difficult from a mathematical perspective, but the unfamiliarity of their content makes them difficult in the context of a popular competition. Problems of this type are perfectly well suited to the MK, but are not categorised as 3-point (i.e., easy) problems.

Putting together all these restrictions, it is perhaps less surprising that it is so hard to find competition problems that are both easy and interesting. There is a very fine line separating routine (textbook) problems and interesting but multiple idea/multiple step problems. By avoiding traps and certain all-too-obvious distractors and ideally sticking to a single content category, further restrictions have to be fulfilled. Considering all of this, there appears to be an inherent contradiction between the requirement that a problem should be ‘interesting’ and the requirement that it should be ‘easy’.

Despite the severely limited pool of problems fitting this description, problem posers are constantly challenged to find appropriate easy and interesting problems. As the annual problem sets confirm, they are often quite successful at overcoming the difficulties.

6 Hard problems in popular competitions

The most obvious reason for a problem to be rated as too hard for the MK involves the level of mathematics required for a solution. If a problem can be solved only by applying some higher-level results that may not be known to all participants, the problem will not be able to engage the non-specialist participant. A typical example of this would be an inequality problem whose proof requires a tool like the general means inequality or the Cauchy inequality. Such tools can be assumed to be at the disposal of students participating in the MO, but they will likely not be familiar to students whose only applicable knowledge is what they have been exposed to in their regular classrooms.

There are also some more subtle reasons for problems to be unsuited for the MK. Sometimes, problems are suggested that seem at first glance to be quite suitable, but in which the result involves a proof whose structure is of more interest than the result itself. Such problems are intrinsically unsuited to the multiple-choice format. A typical example for such a question is the following.

D) Determine the number of triples of integers solving the equation

$$ (6x - 1)^2 + (3y + 1)^2 = z^2. $$

Ignoring for a moment the matter of whether this question is sufficiently interesting for a competition, the fact that there are no such integer solutions (each of the squares on the left side of the equation is congruent 1 modulo 3, while the number on the right side of the equation cannot be congruent 2 modulo 3) can conceivably be of interest only if we ask for the proof of the fact. Suggesting possible answers like

A) 0  B) 1  C) 2  D) 3  E) more than 3,

is beside the point of the mathematical content of the question being posed. The problem may therefore be usable for a competition requiring justification, but not for the multiple-choice environment of the MK.

Then there is the matter of complexity. This can apply to the problem statement itself or the available derivations of correct answers. If the set-up of a competition problem requires several paragraphs to explain, a competitor will require more time to read and comprehend the question than the time that is available. Typically, an MK paper is composed of 30 problems to be solved in no more than 75 min. This allows an average of $2\frac{1}{2}$ min per question. This includes reading the problem, comprehending the given information, solving the problem and entering the answer on the answer sheet. Even allowing for the fact that a student may be done with the easy problems in less than average time, there is still not much time left over for each of the more difficult problems. Any problem setting composed of more than a few sentences (or with a figure requiring lengthy perusal) will simply take too long to be dealt with appropriately.

Similarly, a problem may simply require too many steps of calculation or logical argument than can reasonably be done in a limited time. This does not mean that the problem is in any way inferior, of course. One and the same problem can often be made shorter or longer, depending on the numbers used. An example is the following question, a 5-point problem taken from the Student paper 2019:

(E) Three different numbers are chosen at random from the set \{1, 2, 3, \ldots, 10\}. What is the probability that one of them is the average of the other two?
A) \( \frac{1}{10} \)  (B) \( \frac{1}{6} \)  (C) \( \frac{1}{4} \)  (D) \( \frac{1}{3} \)  (E) \( \frac{1}{2} \)

This is not an easy problem for most students. In order to solve this, they must realize that there are \( \binom{10}{3} = 120 \) different ways to choose 3 numbers from this 10-element set, and that groups of three numbers with the required property must be of the type \( x - a, x \) and \( x + a \). Since the middle number \( x \) can be any of the numbers 2, 3, …, 9, and the number of options for \( a \) for each of these are 1, 2, 3, 4, 4, 3, 2 and 1 respectively, there are \( 1 + 2 + 3 + 4 + 4 + 3 + 2 + 1 = 20 \) such possibilities, and the probability is therefore \( \frac{20}{120} = \frac{1}{6} \). This differs markedly from the following two variations of the problem.

(E1) Three different numbers are chosen at random from the set \{1, 2, 3, 4\}. What is the probability that one of them is the average of the other two?

(E2) Three different numbers are chosen at random from the set \{1, 2, 3, 5, 6, 7, 9, 10, 11, 12\}. What is the probability that one of them is the average of the other two?

These problems require essentially the same reasoning as the original problem E. Still, problem E1 is obviously much easier to answer than E, as there are only four possible choices for the three numbers and listing them all shows us that the required probability is \( \frac{1}{2} \). On the other hand, finding the probability in E2 requires the consideration of many more cases, and is too long for the time allotment of the MK.

With this consideration, it may seem that the avoidance of problems that could be considered ‘too hard’ for the MK should be straightforward, but this is far from the case. There are many subtle factors at play in the problem-selection process that make the distinction between appropriate hard problems and ‘too-hard’ problems quite difficult in practice.

7 Precision is everything. Or is it?

At the MK, a student is faced with many items that are to be dealt with in an average of 2½ min per question. Problem posers must therefore set short and precise tasks, as students will need that time to find their solutions. Unfortunately, a task that is communicated in too abbreviated a fashion may not be clear to the students or, at worst, logically ambiguous. Any attempt at mathematical precision will either require the use of technical terms the students may not have at their disposal, or else the text of the problem will be lengthened. Even worse, this may not even make the problem more understandable. In contrast to MO, however, where students are used to dealing with precise and/or compressed formulations, common notations, and simplifying mathematical symbols, the average participant of the MK cannot be expected to be familiar with this particular ‘language’. The specific vocabulary available for the problem poser also depends on the curricular knowledge of the targeted group of students. In an international group, this will vary from country to country.

The fact that pictures, symbols and signs can be used is of great help for the problem poser, as the following item from Pré-Écolier in 2019 shows:

(F) A mouse and a piece of cheese are in the opposite corners of the board. The mouse can only move as shown by the arrows. In how many ways can the mouse reach the cheese?

(A) 2  (B) 3  (C) 4  (D) 5  (E) 6

Neither the size of the board nor the positions of the two objects nor the possible directions of movement have to be specified due to the available picture and the explanatory addition of the arrows. Such tools make it possible to pack a lot of information into a language-independent format. This is a huge advantage for an international competition, where translation is an issue, but also from the point of view of eliminating the need for unfamiliar technical terms. In the MK 2019, the percentage of items including figures ranges from 37% at the Junior category to 87% at Pré-Écolier. In total, 89 out of a total of 153 items contained figures or pictograms, and unsurprisingly, this number decreases with the increasing age of the participants. Graphic representations are also used as a tool to increase motivation, but to a large extent the intent is for them to replace technical descriptions of mathematical facts and describe geometric objects and operations, such as rotations, reflections, or concrete instances of geometric combinatorics. A further central aspect is that pictograms are used to replace formal mathematical notation and to bypass the use of variables. An example is the following problem from Benjamin 2019:

(G) Bridget folds a square piece of paper twice and subsequently cuts it along the two lines as shown in the picture. How many pieces of paper does she obtain this way?
A language was addressed by Durand-Guerrier (2008). In through the use of arrows and dotted lines makes a crucial difference in decreasing the amount of text required to understand the problem.

While technical terms can often be replaced by figures and pictograms, the targeted use of colloquial language is another option that can be considered for this purpose. However, colloquial language can make the wording of a task either shorter or longer, depending on the task. Under certain conditions, a situation might be described colloquially in an abbreviated, and still understandable manner, as a particular situation need not be mathematically embedded. On the other hand, it might actually make the required text longer, if a mathematical fact is to be expressed with sufficient stringency using everyday language, without resorting to mathematically precise, but obscure, terms. An example of this is the following problem from the Junior paper in 2019, where the mathematical term ‘reflection’ is omitted by embedding in a real-world situation:

(H) A barber wants to write the word SHAVE on a board in such a way that a client looking into the mirror reads the word correctly. How should the barber write it on the board?

(A) SHAVE (B) SHAVE (C) EVAHE (D) EVAHS (E) SHAVE

The problem of possible misinterpretation of colloquial language was addressed by Durand-Guerrier (2008). In the analysis of a particular logic item of the MK 1994 in France, the author stated that the French version of the task can be solved only if a certain specific word is semantically assigned the exact meaning intended by the problem poser. Of course, since problem posers are aware of this source of conflict, they will generally do their utmost to avoid making any formal mistakes, and to simultaneously provide a balance between colloquial language and precision. Often, this is a matter that greatly depends on the specificities of a language. In an international competition like the MK, this means that the greatest attention in this respect must be paid during the translation of the tasks into the individual languages.

The use of colloquial language and sticking to the everyday meaning of words seems to improve the accessibility of problems at a popular competition in general, but care must be taken to avoid ambiguities. Figures and pictograms can describe mathematical situations, paraphrase definitions and often completely replace lengthy formulations. However, when including figures, two main things should be kept in mind: Firstly, forcing young students to switch between figures and text to get the necessary information in order to solve the task, may decrease a pupil’s performance by increasing the cognitive load on the learner’s limited working memory (Berends & van Lieshout, 2009). Secondly, as Carotenuto et al. (2021) pointed out, the variation in the presentation (e.g., by including figures) significantly changes students’ approaches and answers to word problems in the same context, because the elements included in the presentation of a word problem can have a strong informational component for the students. While the first point further limits the creation of adequate formulations of tasks, especially for younger students, being aware of the second point may even lead to greater variability when creating tasks. Summing up, a compromise between short, precise and manifestly attractive tasks is part of the process of creating problems, even if this is easier to do for certain problems than it is for others.

8 Packaging is everything. Or is it?

The internal world of the MK is quite fascinating. In 2019, mice were looking for cheese, mothers cutting cakes, pieces of fruit speaking, kangaroos thinking about their age, giants building sandcastles, barbers attracting customers, ants moving on graphs, and—last, but not least—an endless variation of children dealing with everyday problems like eating chocolate, shaking hands, or throwing balls. On the other hand, some problems simply deal with geometric objects or numbers with certain properties. A problem of the latter, inner-mathematical type, is the following, from the Student paper 2019:

(I) A positive integer \( n \) is called good if its largest divisor (excluding \( n \)) is equal to \( n - 6 \). How many good positive integers are there?

(A) 1 (B) 2 (C) 3 (D) 6 (E) infinitely many

Problem posers often try to use the daily-life experiences of the younger students pragmatically in order to have a contextually broad variation of similar mathematical content. This is necessary, as their mathematical knowledge is still limited. In particular, this can be seen in Pre-Ecolier, where only one of 2019’s tasks is inner-mathematical. At the other end of the age spectrum, there are barely any hard tasks which are not inner-mathematical. An explanation may be the fact that the items are already hard from a mathematical point of view and therefore intentionally stripped of any additional time-consuming complexity that would be
introduced by requiring the translation of some real-world context into mathematics.

Some authors have a great love of packing mathematical problems into stories, literary contexts, and even fairy tales (see Kašuba, 2017). Their argument is that these kinds of problems improve the motivation of the students, especially younger ones. In return, it can sometimes be the case that a context will come across as being somehow artificial or too unrealistic. Studies show a positive effect on motivation in providing real-world and modelling problems, in particular for weak students (see Maaß & Mischo, 2012; Pongsakdi et al., 2019), which is of course a major goal of a popular competition. On the other hand, Rellensmann and Schukajlow (2017) found that ninth grade students experienced the same levels of enjoyment and boredom when solving problems with and without a connection to reality. It should be pointed out that motivation and performance seem to be related: By means of an intervention study, Pongsakdi et al. (2019) found that the problem-solving performance of students in grades four and six, working on a teacher’s innovative self-created challenging word problems, increases when compared to students faced with typical word problems, when the initial motivational level of the students was already high. Furthermore, there is a correlation between text comprehension and performance both for easy and difficult tasks (Plath, 2020; Pongsakdi et al., 2020). Plath (2020) also showed that performance generally decreases when the linguistic complexity of a word problem is increased and the mathematical complexity is simultaneously left unchanged, and that half of the time spent solving a word problem is spent on understanding the setting and stripping it down to its mathematical content.

The research findings stated above can be transferred to the MK only to a certain extent, since that competition contains a very special subset of word problems, namely word problems as exercises in complex problem solving, requiring the use of cognitive strategies (heuristics) as well as metacognitive (or self-regulatory) strategies, due to the use of the classification of word problems by Verschaffel et al. (2020). There does, however, seem to be some evidence suggesting that real-world settings can provide positive motivational effects for many weaker or younger students. On the other hand, it may actually decrease the probability of solving the problem due to the additional challenges posed by the necessity of reading information from additional (and possibly unnecessarily complex) contexts. The question of the extent to which the arithmetical and mathematical problem-solving skills and/or text comprehension should be the main focus seems to have no ultimately correct answer, and is addressed annually by the selection groups. A main challenge for the community is to choose and select a good mix of problems for the MK, where all the various requirements are taken into account.

9 Conclusion

In this paper, we have tried to discuss all relevant properties and aspects of potential competition problems, including the following: differences from textbook problems, content of problems, the crucial role of distractors, the desire for problems to be both easy and interesting, the range of difficulty, the precision and the packaging. Furthermore, we embedded the practitioner’s view of creating problems in research on problem posing and word-problem solving, thereby addressing an important need to connect the Competition Movement and the mathematics education community.

After a close look at some specific problems from the classroom and the world of the competitions in Sect. 2, the de facto competition curriculum was explored in Sect. 3. In Sects. 4–6 the creation of high-quality problems and distractors for the MK was considered, with a special focus on the theory of problem posing. In Sect. 7, details of the formulation of multiple-choice competition problems and the role of figures were discussed. Finally, in Sect. 8, we highlighted the variation of design within the tasks of the MK and identified connections to current research on word-problem solving. This last point could serve as a starting point for more detailed investigations, for instance when focusing on students’ motivation when solving problems of the MK, and it could contribute to current discussions of affective factors in solving complex word problems.

Besides being aware of all these aspects, experience seems to be important in order to create high-quality problems for the MK, as becomes apparent in numerous instances within the paper. Lee (2020) pointed out that “[i]t would be beneficial to conduct problem posing studies with mathematics experts because it can not only provide students with experiences characterizing mathematics experts’ thinking but also assist students to learn the ways of thinking while facing problems” (Lee, 2020, p. 12). Despite the fact that the limitations of the study are quite clear as it derived from the experiences of both authors, we are convinced that our analysis serves as a starting point for a discourse on essential aspects of problems for popular mathematics competitions, as well as for empirically investigating factors that allow rich and productive problem posing for popular competitions. Such investigations could give a better understanding of the nature of problem posing itself and hence enrich the perspective of problem posing as a research goal (Cai & Leikin, 2020).

Ultimately, it must be remembered that the whole point of the exercise is to motivate students and to give them a joyful experience, new insights and a different perspective on the fascinating and versatile world of mathematics. This may sometimes require some difficult decisions, but careful consideration of all these nuances will certainly result
in competitions that participants can ultimately profit from and enjoy.

**Funding** Open access funding provided by University of Graz.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**References**

Akveld, M., Caceres-Duque, L. F., & Geretschläger, R. (2020). Math Kangaroo. Mathematics Competitions, 33(2), 48–66.

Andreeescu, T., Gallian, J. A., Kane, J. M., & Mertz, J. E. (2008). Cross-cultural analysis of students with exceptional talent in mathematical problem solving. Notices of the AMS, 55(10), 1248–1260.

Andritsch, L., Hauke, E., & Kelz, J. (2020). How to create and solve: Analysis of items from the Mathematical Kangaroo from two perspectives. In R. Geretschläger (Ed.), Engaging young students in mathematics through competitions—World perspectives and practices, Vol. II: Mathematics competitions and how they relate to research, teaching and motivation (pp. 117–136). World Scientific.

Applebaum, M., & Gofman, E. (2015). The perception of the concept of a “challenging task” by mathematically promising elementary school students. In F. M. Singer, F. Toader, & C. Voica (Eds.), Electronic Proceedings of the 9th International Conference Mathematical Creativity and Giftedness, Sinaia.

Barzel, B., Leuders, T., Prediger, S., & Hußmann, S. (2013). Designing tasks for engaging students in active knowledge organization. In Watson et al. (Eds.), ICMI Study 22 on Task Design—Proceedings of Study Conference (pp. 285–294). Oxford.

Baumanns, L., & Rott, B. (2021). Rethinking problem-posing situations: A review. Investigations in Mathematics Learning, 13(2), 59–76.

Berends, I. E., & van Lieshout, E. C. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. Learning and Instruction, 19(4), 345–353.

Bray, A., & Tangney, B. (2017). Technology usage in mathematics education research—A systematic review of recent trends. Computers and Education, 114, 255–273.

Bruder, R. (2016). Role of heuristic for problem solving. In P. Liljedahl, M. Santos-Trigo, U. Malaspinia, & R. Bruder (Eds.), Problem solving in mathematics education (pp. 2–6). Springer.

Cai, J., & Leikin, R. (2020). Affect in mathematical problem posing: Conceptualization, advances, and future directions for research. Educational Studies in Mathematics, 105(3), 287–301.

Caretto, G., Di Martino, P., & Lemmi, M. (2021). Students’ suspension of sense making in problem solving. ZDM—Mathematics Education, 53(4), 817–830.

Djukić, D., Janković, V., Matić, I., & Petrović, N. (2011). The IMO compendium: A collection of problems suggested for the international mathematical olympiads: 1959–2009. Springer.

Donner, L., Kelz, J., Stipsits, E., & Stuhlpfarrer, D. (2021). Which test-wiseness based strategies are used by Austrian winners of the Mathematical Kangaroo? Mathematics Competitions, 34(1), 88–101.

Durand-Guerrier, V. (2008). Truth versus validity in mathematical proof. ZDM—The International Journal on Mathematics Education, 40(3), 373–384.

Elgrably, H., & Leikin, R. (2021). Creativity as a function of problem-solving expertise: Posing new problems through investigations. ZDM—Mathematics Education, 53(4), 891–904.

Engel, A. (1998). Problem solving strategies. Springer.

Geraniou, E., & Jankvist, U. T. (2019). Towards a definition of “mathematical digital competency.” Educational Studies in Mathematics, 102(1), 29–45.

Geretschläger, R. (2017). The rainbow of mathematics—Teaching the complete spectrum and the role mathematics competitions can play. In A. Soifer (Ed.), Competitions for young mathematicians—Perspectives from five continents (pp. 145–170). Springer.

Geretschläger, R. (2020). Beyond the rainbow—Thoughts on the potential of mathematics competition problems in the classroom. In R. Geretschläger (Ed.), Engaging young students in mathematics through competitions—World perspectives and practices, Vol. II: Mathematics competitions and how they relate to research, teaching and motivation (pp. 85–116). World Scientific.

Geretschläger, R., Griesel, H., & Postel, H. (2004). Elemente der Mathematik. E. Dorner.

Kašuba, R. (2017). From the lifetime experience of a seasoned math educator—Thoughts, hopes, views and impressions. In A. Soifer (Ed.), Competitions for young mathematicians—Perspectives from five continents (pp. 271–301). Springer.

Kenderov, P. (2006). Competitions and mathematics education. In M. Sanz-Solé, J. Soria, J. L. Varona, & J. Verdera (Eds.), Proceedings of the International Congress of Mathematicians (ICM) (Vol. 3, pp. 1583–1598). Zürich.

Klafkis, M. S. (1994). Mathematical creativity in problem solving and problem proposing II. Mathematics Competitions, 7(2), 39–66.

Koichu, B., & Andžāns, A. (2009). Mathematical creativity and giftedness in out-of-school activities. In R. Leikin, A. Berman, & B. Koichu (Eds.), Creativity in mathematics and the education of gifted students (pp. 285–307). Brill.

Konstantinov, N. N. (1997). Tournir gorodov i matematicheskaya olimpiada [Tournament of the Towns and mathematical Olympiad]. Matematicheskoe Prosveschenie, 3(1), 164–174.

Kontorovich, I. (2015). Why do experts pose problems for mathematics competitions? In C. Bernack-Schüler, R. Erens, T. Leuders, & A. Eichler (Eds.), Views and beliefs in mathematics education (pp. 171–181). Springer.

Kontorovich, I. (2020). Problem-posing triggers or where do mathematics competition problems come from? Educational Studies in Mathematics, 105(3), 389–406.

Kontorovich, I., & Koichu, B. (2016). A case study of an expert problem poser for mathematics competitions. International Journal of Science and Mathematics Education, 14(1), 81–99.

Lee, S. Y. (2020). Research status of mathematical problem posing in mathematics education journals. International Journal of Science and Mathematics Education. https://doi.org/10.1007/s10763-020-10128-z.

Liljedahl, P., & Cai, J. (2021). Empirical research on problem solving and problem posing: A look at the state of the art. ZDM—Mathematics Education, 53(4), 723–735.

Millman, J., Bishop, C. H., & Ebel, R. (1965). An analysis of test-wiseness. Educational and Psychological Measurement, 25(3), 707–726.

Maas, K., & Mischo, C. (2012). Förderd mathematisches Modellieren die Motivation in Mathematik? Befunde einer Interventionstudie bei HauptschülerInnen. Mathematica Didactica, 35, 25–49.
Plath, J. (2020). Verstehensprozesse bei der Bearbeitung realitätsbezogener Mathematikaufgaben: Klassische Textaufgaben vs. Zeitungstexte. *Journal Für Mathematik-Didaktik, 41*, 237–266.

Pongsakdi, N., Kajamies, A., Veermans, K., Lertola, K., Vauras, M., & Lehtinen, E. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM—Mathematics Education, 52*(1), 33–44.

Pongsakdi, N., Laakkonen, E., Laine, T., Veermans, K., Hannula-Sormunen, M. M., & Lehtinen, E. (2019). The role of beliefs and motivational variables in enhancing word problem solving. *Scandinavian Journal of Educational Research, 63*(2), 179–197.

Poulos, A. (2017). A research on the creation of problems for mathematical competitions. *Teaching of Mathematics, 20*(1), 26–36.

Rellensmann, J., & Schukajlow, S. (2017). Does students’ interest in a mathematical problem depend on the problem’s connection to reality? An analysis of students’ interest and pre-service teachers’ judgments of students’ interest in problems with and without a connection to reality. *ZDM—Mathematics Education, 49*(3), 367–378.

Reznick, B. (1994). Some thoughts on writing for the Putnam. In A. H. Schoenfeld (Ed.), *Mathematical thinking and problem solving* (pp. 19–29). Routledge.

Sharich, V. (2017). Goals and challenges of the mathematical Olympiads of today: Science, sport, university admission, or status? In A. Karp (Ed.), *Current issues in mathematics education* (pp. 67–78). Bedford.

Sharigin, I. F. (1991). Otkuda berutsia zadachi? [Where do problems come from?] Part I, *Kvant, 8*, 42–48; part II, *Kvant, 9*, 42–49.

Shi, H. (2012). Helping solve mathematics competition problems using computer technology. *International Journal of Advanced Research in Computer Science, 3*(4), 23–26.

Sinclair, N., Arzarello, F., Gaisman, M. T., Lozano, M. D., Dagiene, V., Behrooz, E., & Jackiw, N. (2009). Implementing digital technologies at a national scale. In C. Hoyles & J.-B. Lagrange (Eds.), *Mathematics education and technology-rethinking the terrain* (pp. 61–78). Springer.

Smith, J. K. (1982). Converging on correct answers: A peculiarity of multiple choice items. *Journal of Educational Measurement, 19*(3), 211–220.

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM—Mathematics Education, 52*(1), 1–16.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.