

Fairness

“That’s not fair!”

Whenever someone sees something happening at a tournament that they don’t like, there’s a good chance that the complaint, whatever its origin or merit, will be couched in terms of fairness. Usually, in fact, it’s not fairness that’s the issue, but its opposite: unfairness.

Fairness is crucially important in tournament design. If the players see a tournament as unfair, it’s likely in for trouble, no matter what its other virtues. Fairness is not the only legitimate goal of the tournament director—often some measure of fairness must be sacrificed in favor of one of the other objectives. But fairness is important enough that it should only be compromised in aid of some important goal. A tournament should not be gratuitously unfair.

It is of the essence of good tournament practice that the organizers should be able to give a sensible response to fairness complaints. And that begins with thinking clearly about the concept of fairness itself. Many times, the dispute behind the cry of “that’s not fair” has no clear solution. The problem is not that people don’t know what fairness is or how to achieve it, but that they are talking about different kinds of fairness. Fairness is not a unitary concept. The first step in making fairness questions tractable is to deconstruct the concept of fairness.

The word *fair*, in the context of tournaments, often seems to be used so generally that it’s become a synonym for *good*. Anything that’s bad about a tourney is unfair. And, indeed, the word *fair* in general use has meanings so various that it can be made to stand for almost any positive attribute. The Oxford English Dictionary devotes six columns to explaining 18 different senses of the word *fair* (and doesn’t include one of the ones I think is particularly relevant to tournaments). For our purposes, however, it makes sense to relieve *fairness* of some of this burden, and other tournament virtues are discussed above under the headings of participation, spectacle, and efficiency.

Even after offloading some of the possible meanings of *fairness*, however, the term remains ambiguous. There are three different kinds of fairness. And these three sorts of fairness are different enough that they are often in conflict with each other. An action that increases fairness in one of its meanings often diminishes fairness in another. Many questions of fairness are unresolvable because they are addressed to different aspects of fairness.

Here are the three kinds of fairness:

Fairness (A) is the fairness of met expectations. When a tournament isn’t run the way people expect it to run, they often feel that it’s not fair.

Fairness (B) is the fairness of equality. When not everyone gets an equal chance to win, a tournament is often judged unfair.

Fairness (C) is the fairness of meritocracy. People expect tournaments to be run in such a way that the winners are the people who play the best. Features of a tournament that hinder this from happening are often thought unfair.

Most people have a more synthetic concept of fairness that includes some measure of each of the three kinds. A narrow focus on one sort of fairness will lead to ridiculous results.

If, for example, fairness (B) was the only thing that mattered, you could run a perfectly fair tournament without the trouble of actually playing any games. You could simply have a drawing for the trophy and the prize fund, which would treat every player equitably. Perfectly fair in the fairness (B) sense, but entirely unfair in terms of fairness (C).

If, instead, you care only about fairness (C) you might also be able to dispense with playing games. If one of the players is known to be better than any of the others, you can simply hand that player the trophy and the money. This way, you're being perfectly fair in terms of fairness (C), and perfectly unfair in terms of fairness (B).

Almost no one would consider either of these designs a fair tournament, because almost everyone has a notion of fairness that include elements of both fairness (B) and fairness (C). In fact, almost everyone's notion of fairness also has, at least in most contexts, a generous dollop of fairness (A) as well.

The key to resolving questions of fairness begins by taking the concept of fairness apart into its three components and asking which kind of fairness is at stake. If a dispute arises because the two types of fairness are in conflict with each other, and the two sides don't agree on which type should have priority, it may be irreconcilable. But even in that case, clear thinking about the kinds of fairness involved can help make the disagreement less toxic. It's one thing to say, "you're being unfair", and quite another to say, "you want to do it that way because you value tradition (i.e., fairness (A)) more than I do."

When it happens that the two sides in a fairness dispute are talking about the same kind of fairness, the dispute may have a clear answer. One of the chief aspirations of tourneygeek is to provide such answers when they're available.

To understand which fairness questions have answers and which do not, we'll begin by discussing each of the three kinds of fairness separately.

Fairness (A)

Fairness (A) is the fairness of meeting expectations. Regardless of whether a particular practice is equitable (appealing to fairness (B)), or meritocratic (fairness (C)), tourneys are often judged to be unfair because they're not run the way that people expect them to be run.

Most of the discussion of fairness in this book concerns fairness (B) or, especially, fairness (C). There's a good reason for that—we seek, where we can, to provide clear answers to questions or whether something is more or less fair, and so tend to concentrate on the two sorts of fairness that are, as we'll show presently, at least somewhat quantifiable. Fairness (A), in contrast, is unquantifiable because people's expectations tend to be qualitative rather than quantitative. Disputes about fairness (A) can rarely be definitively resolved.

But this does not mean that fairness (A) is unimportant, or that there's nothing sensible to say about why some fairness (A) claims are stronger than others.

The strength or weakness of a fairness (A) argument depends on two factors: the source of the expectation, and the degree to which that expectation is relied upon. Considering the source goes to whether the expectation is reasonable, and considering the degree of reliance goes to whether the expectation is consequential.

The Source of the Expectation

Here are a number of possible sources for fairness (A) expectations, in roughly descending order of strength:

1. The rules of the game being played.

A game is defined by its rules, and a tourney that does not obey those rules is always suspect. But even in this case, expectations can differ. Many games have core rules that are considered more essential than others. Often some of the other rules of a game are set out in some official rule set that is known only to very serious practitioners. The highly technical rules, for example, governing what constitutes a legal serve in table tennis are expected to be honored at a professional-level tournament, but not necessarily at the tourney you organize at a church picnic.

2. Announcements from tournament organizers.

Within the rules, tournament organizers often have considerable latitude to structure the competition. But they are expected to announce in advance what their choices have been, and to honor those announcements. In order to deviate, there need to be unforeseen circumstances that make the announced practices clearly undesirable, and even then most organizers will try to stick to their statements unless it is impossible to honor them.

3. Written policy of a governing body.

If a competition acknowledges the authority of some governing body, that body will often have ancillary rules that govern the ways that tournaments are run. In some cases these are very detailed—a tennis tournament run under the auspices of the Association of Tennis Professionals can be expected to run in conformity with a 400-page official rulebook (which does not cover such mundane matters as the rules of tennis).

Whether or not a particular competition is governed by a particular body is not always clear. Some organizations claim jurisdiction over the game itself, and gird themselves with as much apparent authority as they can muster. For example, the self-appointed guardian of the rules of golf in most of the world calls itself the R&A, presumably to encourage us to think that, despite the fact that it was created in 2004, it is both royal and ancient.

4. Conformity with past tournaments in the same series.

When a tourney is a part of a series, it makes sense that it should, unless otherwise specified, be run the same way. The strength of this expectation depends on how tightly integrated the series is seen to be, and strongest when the series can be considered an event in its own right. When an event is an annual event, the presumption that it will be run the same way as it was run last year is stronger when the event is well established, and relatively weak for any “second annual” event.

5. Conformity with similar tournaments for the same activity.

For many games and sports there comes to be a general understanding of how tournaments should be organized. People who have been to a lot of tournaments reasonably expect that tourneys will be run in accord with that general understanding.

For example, most professional tennis tournaments use a distinctive tiered seeding that's uncommon for most other sports. This is a matter of written policy for tennis tournaments governed by the rules of the ATP or the WTA, but because it is established practice for those highly regarded events it is also considered fair for tennis tournaments in general. But for other sports, which have no tradition of tiered seeding, the same practice is likely to be considered unfair.

6. Practice in similar tournaments.

Finally, there are some expectations that are founded on a generalized understanding of how tournaments in general should be run. A player who has no expectations for how the tournaments are run specifically for a particular game may still have an expectation based on other games they've competed in.

If you search, for example, for brackets to use in a double-elimination tournament, you are quite likely, unless you stumble upon this book or a site like *tourneygeek*, to find only unshifted lower brackets. This gives rise to a generalized feeling that there's something wrong about a shifted double-elimination bracket.

Reliance on the Expectation

Whatever the source of an expectation, a fairness claim is much strengthened when the claimant can show that they've reasonably relied upon that expectation.

Perhaps the most common kind of reliance is based on a player's willingness to enter the event in question. A player has a stronger fairness (A) objection if they can credibly make a claim like one of these:

I wouldn't have entered if I'd known that ...

... I wasn't guaranteed at least three matches;

... expert players would also be allowed to compete in this "intermediate" event;

... the tournament was only returning half of the entry fees in prizes;

... my favorite (putter, bidding convention, or whatever) wasn't allowed;

... the funds raised were going to (Planned Parenthood/the National Rifle Association).

The fairness claim is especially strong when the player can show that their reasonable expectation led them to play in a certain way. This is most likely to happen when they're called upon to allocate their resources between individual matches based on their understanding of what other matches they may need to play.

For example, in a double-elimination baseball tournament, teams might handle their pitching staff quite differently depending on whether or not they think that there will be a recharge round. A team that's saved its best pitcher to start the recharge game will be very unhappy to discover that there will be no such game.

Fairness (A) as Tradition

The vaguest sort of fairness (A) claim is one that is based on a sense of tradition. And some competitions (and some people) have a stronger attachment to tradition than others.

To some extent, tradition accrues naturally to a game or sport that has been played for many years. Among the many modern sports that trace their lineage to an origin in the British Isles, two with, arguably, the longest histories are golf and cricket. The earliest known rule books for both sports date to 1744. Both of these sports are generally regarded as more encrusted with tradition than newer sports such as rollerboarding and ultimate. A respect for tradition is inculcated in new players along with specific game skills. It's reasonable to expect fairness (A) claims to be more common and more strongly asserted in more traditional sports, which is sometimes a problem when these sports attempt to evolve to suit modern conditions.

It is not an accident that the body principally in charge of renovating the rules of golf for modern conditions in most of the world calls itself the R&A, echoing the name of the Royal and Ancient Golf Club of St. Andrews. The R&A was created in 2004, but it knows its constituents. There are, no doubt, some left-wing activists who enjoy playing golf, but the organization knows that golfers in general tend to be the sort of people for whom *royal* and *ancient* are positive qualities.

A sport need not be ancient in order to claim meaningful traditions. Ultimate was invented (as *ultimate frisbee*) in 1968, and yet emphasizes its tradition. Under the rubric of *spirit of the game*, ultimate embraces notions of sportsmanship very similar to those traditionally associated with cricket and golf.

Traditions are often much less ancient than commonly imagined. It is widely thought, for example, that Scotsmen have been wearing kilts in their distinctive clan tartans from time immemorial. Not so:

... the kilt is a purely modern costume, first designed, and first worn, by an English industrialist, and that it was bestowed by him on the Highlanders in order not to preserve their traditional way of life but to ease its transformation: to bring them out of the heather and into the factory.¹

The distinctive tartans of the various clans were assigned to them more or less arbitrarily in the early 19th century.

When considering a fairness claim based on tradition, then, it is often a good idea to ask how well established the tradition really is. Often, what is regarded as a tradition is quite a recent innovation.

¹ Hugh Trevor-Roper, "The invention of tradition: the highland tradition of Scotland", in *The Invention of Tradition*, Hobsbawm & Ranger, eds, Cambridge U. Press (1983).

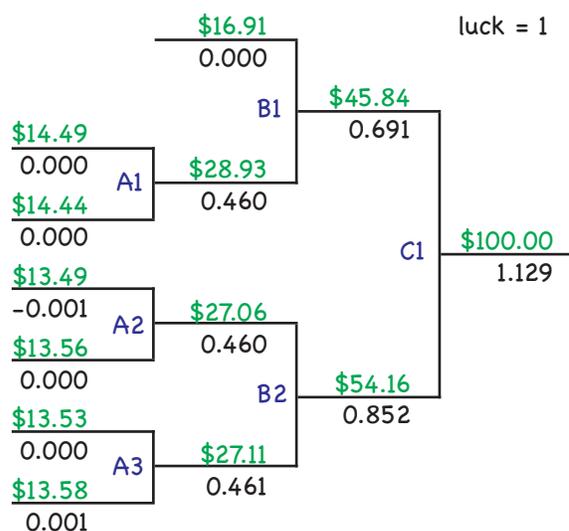
Fairness (B)

Fairness (B) is the fairness of equal opportunity. It is the kind of fairness that springs most easily to many people's minds. It is often the easiest to achieve. But, oddly, it is also the only one of the three types of fairness that is often sacrificed entirely.

Equality is a mathematical notion, so fairness (B) is relatively easy to express in numerical terms. In this section, I'll introduce a numerical measure that will give more substance to the general notion of fairness (B) by allowing levels of inequality to be measured.

To see fairness (B) in action, let's return to the simple, eight-team single elimination bracket discussed in chapter 1, except this time let's make it a seven-team tourney. That means that one team will get a bye in the opening round.

At the right is the analyzed bracket. Now that there's a bye in the first round, the starting lines are no longer about equal. When the bracket was full, all of the starting lines in the A round were worth \$12.50. Now that there's one fewer player, the \$12.50 that would have been expected by that player goes to others. But it hasn't been spread evenly. The lucky team that draws the bye gets almost \$4.50 extra. The two teams that play A2 each get about \$2. This is



\$	16.91	1.184
\$	14.49	1.014
\$	14.44	1.011
\$	13.49	0.944
\$	13.56	0.949
\$	13.53	0.947
\$	13.58	0.951
	100	8.66

because, if they win, they'll draw the unproven bye team in the next round, and so have a better chance to win. The teams in the lower half of the bracket benefit to a lesser extent, each getting about \$1.

Here's a fairness (B) problem. Not everyone has an equal chance.

It's useful to have a way to compare fairness (B) problems numerically, so we define a statistic. To calculate the fairness (B) measure for this bracket, you normalize the expectations for each entry line by dividing it by the mean expectation. The fairness (B) measure is simply the standard deviation of those normalized expectations, multiplied by 100. The result for this bracket is 8.66.

Now, to some extent the inequality of outcomes comes from the random noise—recall that even a million trials was not enough to yield a perfect \$12.50 expectation for every entry line in the full bracket. The fairness (B) measure for the luck = 1 bracket on page x is 0.36—not high, but not, as it should be, zero.

One reason that the fairness (B) measure is multiplied by 100 is to take advantage of an observation I've made when running simulations of brackets that, like the full 8 bracket, are structurally fair from a fairness (B) perspective. That is that when there are enough iterations on

a bracket to yield usable results for other purposes, the measured fairness (B) of an equitable bracket is less than one. As a practical matter, then, I consider any measured fairness (B) of less than one to be a fair bracket, and the measured unfairness the result of random variation.

It's important to note that what's been defined here might more accurately be called a measure of unfairness. Lower numbers are better.

Fairness (B) can be calculated for individual rounds also, and it will frequently be helpful to do so. The number that's calculated for the bracket as a whole is based on the entry lines, in whatever round they occur. In this instance, six of the entry lines are in round A, and one in round B. For round A by itself, the calculation yields 3.44. So round A is not a particularly inequitable round for the six players who play there—the inequality comes because of the one player who skips round A altogether. But for the B round the inequality is much worse, yielding a figure of 21.67. It settles down, a bit, for the C round, which comes in at 11.67.

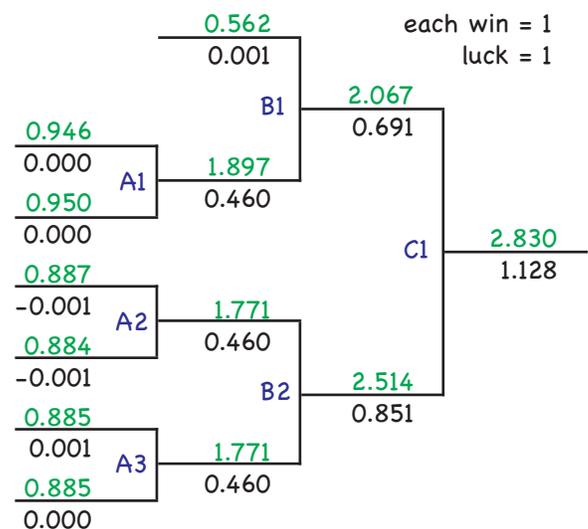
To keep the various forms of the fairness (B) statistic in line, fairness (B) statistics are reported with the following convention. A fairness (B) calculation that's based on entry lines, and thus is characteristic of the bracket as a whole is simply called fairness (B). Where the calculation is based on a particular round it's reported with a lower-case "b", a colon, and the round identifier, for example fairness (b:A). Frequently, all of the entry lines in a bracket are in the A round, in which case fairness (B) and fairness (b:A) are the same thing. Especially where round results are involved, "Fairness" is sometimes abbreviated to "F", to yield something like this: F(b:A).

The example we've considered so far was for a high-skill, blind-draw tourney, with a single \$100 payout to the winner. Changing any of these things will affect the measured level of fairness (B).

If the same bracket is played for a high-luck game, fairness (B) soars to 24.93. When skill is high, the one player drawing the bye faces stiff competition in the B round because of the steep skill progression. With more luck, and hence less skill progression, both the B and C rounds are easier. While the overall level of fairness (B) increases markedly, the individual rounds are fairer: fairness (b:A) = 1.46; fairness (b:B) = 7.21; and fairness (b:C) = 3.65.

Because the fairness (B) calculations are based on the distribution of rewards in the tourney, they can change markedly when the payout scheme is changed. Let's suppose that our seven-team tourney has no prize fund to be distributed, and that the only benefit to the players is the psychic reward they get from winning individual matches. Since there are six matches to be played, the total rewards for the players sum to six.

Here what was the great good fortune of the player who draws the bye turns into a stiff penalty. That player loses the chance of a win in the A round, but still faces much stiffer competition in the B and C rounds due to the skill progression. The



result is much less equitable, with fairness (B) = 15.57. $F(b:A) = 3.58$; $F(b:B) = 41.88$; $F(b:C) = 13.80$.

All of these fairness (B) problems flow from the presence of the single bye. In general, most fairness (B) problems are associated, in one way or another, with byes. The chapter on byes will discuss these issues in much greater depth.

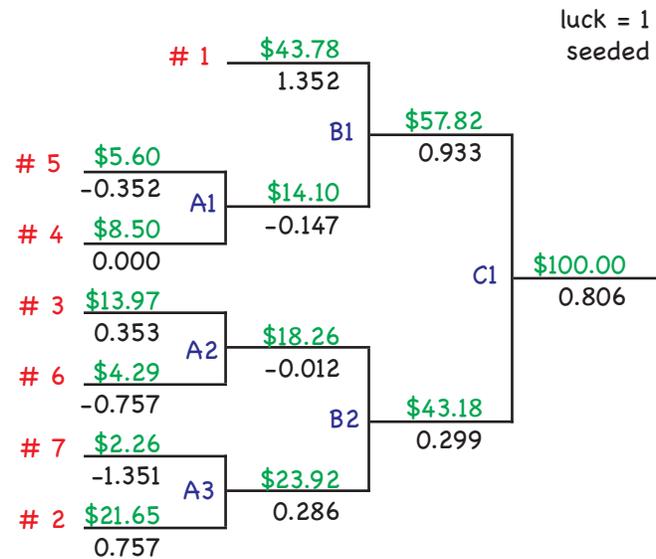
But severe fairness (B) issues also arise due to seeding. When you decide to seed a tournament, you're pretty much abandoning any pretense of valuing fairness (B). In most cases, instead of trying to draw an equitable bracket, you're drawing a bracket that's been purposely manipulated to aid the better players at the expense of the less skillful ones.

The hows and whys of seeding are discussed in much greater detail in chapter X. But as a preview, we'll look at our seven-team bracket seeded in the most conventional way. The bye goes to the top-seeded player, and is now considered an earned bye.

Fairness (B) soars to 102.10. $F(b:A) = 77.38$; $F(b:B) = 52.53$; and $F(b:C) = 20.50$.

Fairness (B) numbers are rarely reported for seeded brackets because it is assumed that no one cares. Seeding is almost always in derogation of fairness (B).

This is not to say that tournament directors should entirely disregard considerations of fairness (B) when running seeded tournaments. In all matters other than loading the initial bracket, inequities should be avoided.



Fairness (C)

Fairness (C) is the fairness of meritocracy. If fairness (A) is giving people what they expect, fairness (C) is giving them what they deserve. In essence, fairness (C) represents the idea that in a tournament the prizes should go to the people who play the best game.

Fairness (C) is the kind of fairness that is easiest to overlook in the context of everyday life. I once gave a talk on fairness to a group of about 20 volunteer mediators. Before I began, we went around the room, and each person said what “fairness” meant to them. The definitions were about evenly split between fairness (A) notions of meeting expectations and fairness (B) ideas of treating everyone alike. But not a single person mentioned anything that sounded much like fairness (C).

There are good reasons why we avoid talking about fairness (C) in everyday life. Fairness (C) implies value judgments that are often difficult and uncomfortable. In order to assess whether Joe has received his fairness (C) due, I first have to decide how good a person Joe is, and whether he is better or worse than someone else who might claim whatever benefit is being dispensed.

Perhaps Joe is a superior human being who is richly deserving of the great wealth and high social standing that he enjoys. But then perhaps he’s really an unscrupulous person who has cheated others of their due, or exploited flaws in “the system” to reap unjust rewards. Perhaps Joe has just been lucky. Perhaps all three of these explanations have some validity in explaining Joe’s success.

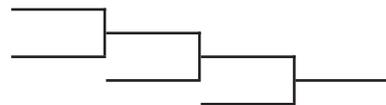
But if the causes of success in life are difficult to discern, we expect something more definite in the protected little slices of life we call tournaments. Whether or not we think Joe is a good human being, if Joe has won a lot of trophies at chess tournaments, we’d like to think that he must be a superior chess player.

The goal of the tournament administrator is to strengthen the relationship between merit and results. We want to eliminate, as thoroughly as we can, the possibility that a positive result comes from cheating, or from exploiting flaws in the system. After that, we usually want to maximize the positive effect of good play, and to minimize the effect of good luck.

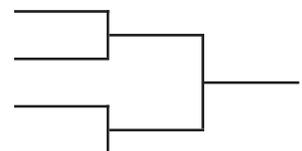
Fairness (C) and Luck

Fairness (C) has an inverse relationship with luck. The more the outcome of an individual match is determined by luck, the less it is determined by skill.

If there is no element of luck, so that the best player always wins any individual game, the task of designing a tournament would be akin to simple sorting. You could reliably identify the best player in a group with a one-pass bubble sort, starting with any two players, and then bringing in new players one by one in subsequent rounds, with the winner of the last game facing the next. In tournament terms, this would be a cascade bracket.



The cascade bracket does have a legitimate use, which will be discussed elsewhere. But for most purposes it is about the least fair bracket imaginable. The tournament designer knows that you can run a fairer

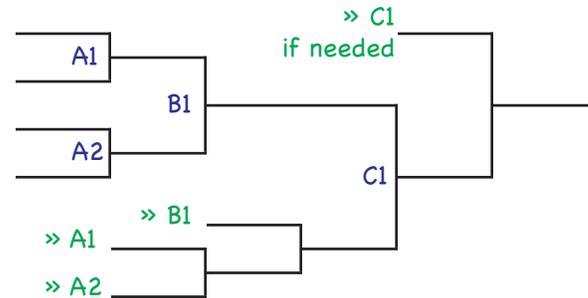


tourney with the same number of games by stacking them differently, as a binary tree rather than as a cascade.

This much you can accomplish simply by noticing that that cascade bracket is not fair from a fairness (B) perspective, because the path to victory is longer for some entrants than for others.

But there is still a fairness problem with the binary tree bracket, in that in any of the three matches, the better team can, with bad luck, lose to the weaker one, and the bracket has no way to allow the better team another chance to prove its mettle.

Now, there are two good ways to improve the fairness of this bracket. The first method is simply to make each match a best-two-out-of-three contest rather than a single game. The second way is to make the tournament into a double-elimination contest, adding matches to create a bracket like this one. (The “if needed game is called a recharge round, and is played only when the winner of B1 loses match C1.)



Either of these approaches gives a player who loses a single game a way back to winning the tournament as a whole. Setting aside, for the moment, the obvious implications for the practical running of the tourney, which is better from a fairness (C) standpoint?

The answer, even in a tiny bracket like this one, is it depends. (You need to get used to this answer, because it will return over and over again.)

To answer the question, first we need a way of measuring fairness (C). I'll show how fairness (C) is calculated by working through an example with the simple binary tree bracket shown above.

Let's say that for one particular iteration of this tourney, the four players draw these Z scores to define their skill levels: 0.892; 0.810; 0.155; and -0.506.

The first step is to find the ideal result—the result expected when the best player does, in fact, win the tourney.

Start by sorting players by their skill level, from highest to lowest. Then sort the available prizes, from highest to lowest, and put these in a parallel column. Multiply the first column by the second column, and sum the results. This score is the ideal payout, the one you get when the fairest possible result is reached. Add another column that shows the prize actually awarded by position for a given run, multiply that by skill, and sum to get an aggregate actual payout.

Fairness (C) is simply the difference between the ideal payout and the actual payout.

Here's how the calculation works:

For this single run of the tournament, the winner was the second-best player. But the skill difference between the best player and the one who won is small so the fairness (C) statistic for this particular run is small, only 8.20. When the top player wins, the fairness (C) statistic is zero. When the third-best player wins, it is 73.7, and on those rare occasions when the worst player wins, it is a whopping 139.8.

rank	Skill	Ideal		Actual	
1	0.892	100	89.2	0	0
2	0.810	0	0	100	81.0
3	0.155	0	0	0	0
4	-0.506	0	0	0	0
			89.2		81.0
		Fairness (C)		8.20	

Running this tournament in the simulator 100,000 times (with luck = 1), I find that the best player wins about 47.2% of the time, the second best 40.6%, the third 10.7%, and the fourth 1.5%. On average, then, the fairness (C) is about 13.30.

This is not, however, a fairness (C) value that's characteristic for the bracket as a whole, only for the bracket when it has players of those four skill levels. As it happens, those skill levels are conducive to fair results because there is such a small skill differential between the top two players. Running another 100,000 trials, but this time drawing four new random skill levels for each trial, the fairness (C) statistic is about 17.52.

When the luck factor is greater, the fairness (C) statistic will rise. The same four outcomes yield the same four figures for fairness (C) for any individual run. But the less fair results will happen more often, and so the fairness (C) average will balloon. At luck = 3, the worst player wins about eight times more often. The average fairness (C) for these three players is 34.87, and for random skill levels, it soars to 55.83.

The calculation is only a little more complicated where there's more than one prize involved. Here, instead of giving 100% of the prize fund to the champion, let's split it 65/35 between the champion and the runner up. Here's the spreadsheet for a single run where the champion was the third player, and the runner up the first:

rank	Skill	Ideal		Actual	
1	0.892	65	57.98	35	31.22
2	0.810	35	28.35	0	0
3	0.155	0	0	65	10.08
4	-0.506	0	0	0	0
			86.33		41.30
		Fairness (C)		45.04	

With both a winner and a runner-up to be paid, there will be 12 different possible results rather than just four. Again, redrawing four new players for each trial, the fairness (C) figures are 19.45 for luck = 1, and for luck = 3 it is 46.17.

Now that we've got a fairness (C) measure to work with, let's return to the question of which of our little four-player tourney designs is better. First, simulation results where there is a winner-takes-all payout:

Luck = 1, winner takes all			Luck = 3, winner takes all		
single tree	2/3 tree	double elim.	single tree	2/3 tree	double elim.
17.52	9.40	9.84	55.83	40.39	42.96

Both designs improve substantially on the simple tree, but at either luck level there's more gain in fairness from playing best two-out-of-three than there is from playing the double elimination. So does this settle the question?

No, because there's another factor we haven't considered yet, the payout scheme. If, instead of giving 100% of the prize fund to the champion, we divide it into two prizes: 65% for first place, and 35% for second. Then the results look like this:

Luck = 1, payout 65%/35%			Luck = 3, payout 65%/35%		
single tree	2/3 tree	double elim.	single tree	2/3 tree	double elim.
19.45	13.55	9.41	46.17	35.90	35.74

This table shows that if you care about second place, the situation is different. With a high level of skill, the double elimination improves fairness (C) a bit, but the result for playing best two out of three is worse than it is for the winner-takes-all situation. This makes sense. One third of the time, the two best players will be drawn into the same side of the bracket, and one of them will not cash.

But if you're running a high-luck event, the result is not so clear. The fairness (C) statistics are so close together that it takes 1,000,000 trials to find a significant difference. For practical purposes, this a dead heat.

Limitations of Fairness (C)

The ability to measure fairness (C) is perhaps the most powerful tool in the workshop of those who seek to provide a reasoned answer rather to many issues in tournament design that are considered, by others, simply matters of tradition or personal preference. It's the fondest hope of tourneygeek that careful study will lead not just to some changes in the way tournaments are run, but to changes that can be considered progress rather than mere fashion.

But if fairness (C) is a powerful tool for understanding tournaments, it is also rather a dangerous one. It is easily capable of leading us into error rather than to knowledge when used incorrectly. It is crucial to be as aware of its main features, and the limitations that those features impose on its application.

It has already been shown that fairness (C) has an inverse relationship with the luck factor of the game involved. There is a reliable relationship here: more luck will invariably cause the

fairness (C) number to be higher. But this increase will happen at different rates for different designs, and so a conclusion based on the simulation at one luck level will not necessarily hold for another. Designs that serve a high-skill game, like tennis, will not necessarily work as well for a high-luck game, like baseball.

Like the luck factor, fairness (C) is also heavily influenced by the number of players or teams involved. As a general rule, the more teams there are, the more opportunities there are for the allocation of prizes to be sub-optimal, and the higher the measured value of fairness (C).

This influence is strong enough to overcome other important fairness (C) influences. For example, we know intuitively that a tourney with a full 16 bracket is fairer than one where there are only 15 entrants, and hence one bye. Byes are a source of unfairness in themselves. But that doesn't mean that the measured fairness (C) for the 15-player tourney will be worse than that for the full bracket. It's simply not appropriate to compare numbers from tourneys with different numbers of entrants.

It's not just the number of entrants that affects fairness (C), but also the way they're chosen. In almost all of the simulations reported in this book, fairness levels are drawn from a Gaussian distribution. But sometimes a particular tourney seems to call for some part of that distribution to be excluded. In particular, an elite tourney draws entrants only from the upper end. This is all very well if it more closely approximates some particular tourney, but it does mean that the fairness (C) results will not be comparable to tourneys without an elite threshold. In general, limiting the range from which entrant skills are drawn has the effect of increasing the effect of luck because the skill levels will be closer together.

Perhaps the most severe limitation on the application of fairness (C) is that it cannot be used to compare tournaments with different payout schemes. It would be lovely to use fairness (C) to compare the fairness of different payout schemes, but experience shows that it just doesn't work.

In sort of the same way that fairness (C) will favor tourneys with fewer entrants, it also favors payout schemes that pay more places, spreading the wealth more widely. When you get to the extreme where each entrant receives the same prize without regard to performance, the measured fairness (C) will be zero—the tourney never rewards a player unjustly, but it also never rewards true merit.

This inability to compare different payout schemes can make it impossible to compare designs. In the example shown above, we were able to compare the single-elimination and double-elimination brackets because both of them would work with either a winner-takes-all payout or a 65/35 split between the top two places.

It would have been nice also to compare these two brackets with a Swiss system tourney, or with a round robin. But it can't be done because those neither of those tourneys, at least in their pure form, make sense with both a winner-takes-all payout and a 65/35 split payout. The reason for this will be discussed below, in the chapters for Swiss and round robin tourneys.

Some tourneys are played with no prize funds. It's possible to use fairness (C) in such circumstances, but only if you can find a way to represent whatever other value it is that players derive from your tourney. In a recreational sports league, for example, you might find that the chief reward for winning one round is simply the opportunity to play another round. If that's the

case, you need to use a payout schedule where every team gets paid according to the number of games it plays.

In any case, fairness (C) will give misleading results when the payouts you assume do not track the values of the people who play the tourney. And it is often a mistake to use a winner-takes-all payout unless it's really true that the only thing you care about is who won.

You can't, for example, use a winner-takes-all payout to assess any tourney that includes a consolation bracket or a last chance bracket because that would mean that fairness (C) is entirely unaffected by what happens in those brackets. But even when there's still some chance that a player can fight through a lower bracket to win the tourney, that chance is often so small that the effect on fairness (C) of poor designs will be negligible. Let's say, for example, you're running a double-elimination. If you're paying only the overall winner in round N, it may be hard to spot the effect of a bad drop into round F, but you're much more likely to be able to see it if there's some payout for reaching round G.

An alarming example of the danger of using winner-takes-all payouts is the phenomenon of the *ugly bottom*. In some rather peculiar double-elimination formats, the lower bracket is so poorly constructed that the winner of the lower tends to be a poor team. And yet these badly-constructed lower brackets sometimes produce fairness (C) statistics that are comparable, or even a little better, than well-constructed lower brackets. That's because by tending to produce weak opponents for the winner's bracket winner, it diminishes the chance of some upsets that would tend to damage fairness (C). But it's hardly a real advantage for a tourney to score well by choosing one pretty good player, and then making everyone else look bad. Paying even one more place will expose the ugly bottom—the system can't look good by producing a weak second-place contender if it's got to pay second place.