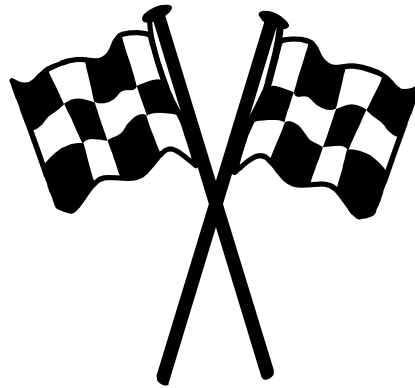


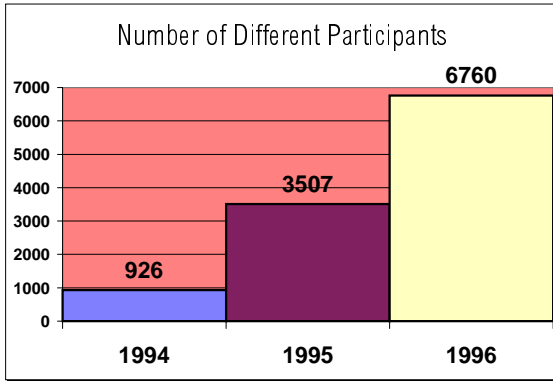
MORE THAN ENOUGH ANSWERS TO
QUESTIONS
NOBODY ASKED ABOUT
F1P6
ick

(Formula 1 Grand Prix Pick 6 competition)



- ◇ *What does the growth pattern look like?*
- ◇ *Are the guesses getting any better?*
- ◇ *How fair is the scoring system?*
- ◇ *Who is the best picker in the last three years?*
- ◇ *Is the whole smarter than the average of the parts?*
- ◇ *Are there horses for courses, or should you let it ride?*

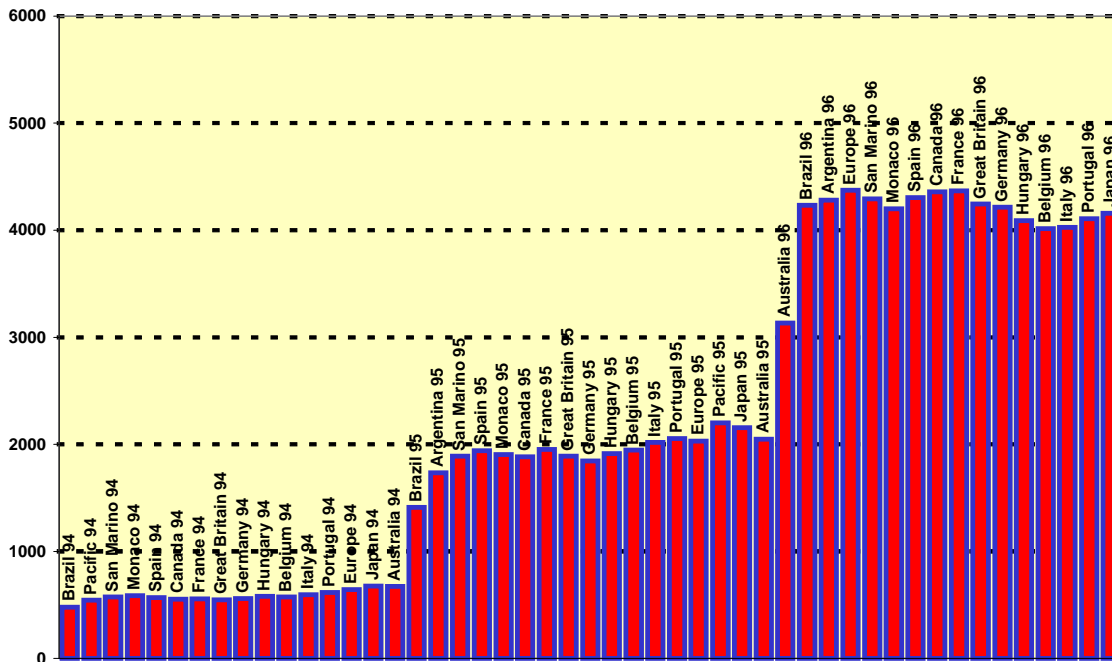




Growth...

The number of people submitting entries to the Pick 6 competition has grown 600% in two years. The chart below shows that the growth occurs between years; the number of entries per race is stable within each year.

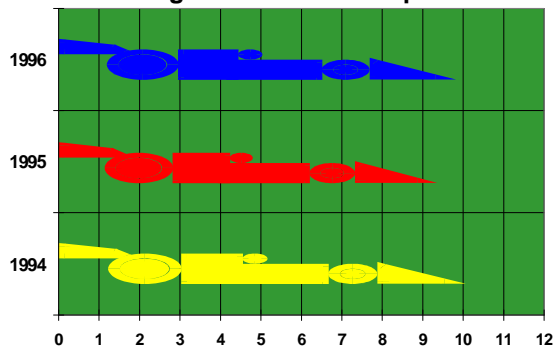
Number of Participants In Each Grand Prix



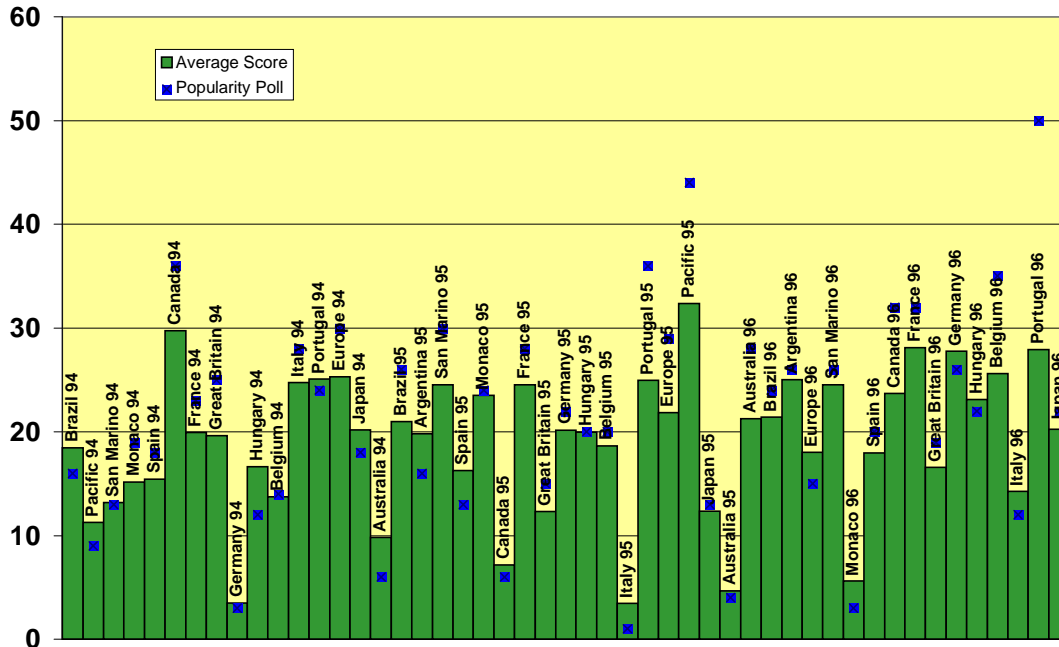
Persistence...

Most SixPickers don't enter every time. Even though entries are carried over once to the next race if no new entry is received, the average number of races entered has consistently been between 9 and 10 each season. Strangely, 1995, with one more race, had the lowest average number of entries.

Average Entries Per Participant



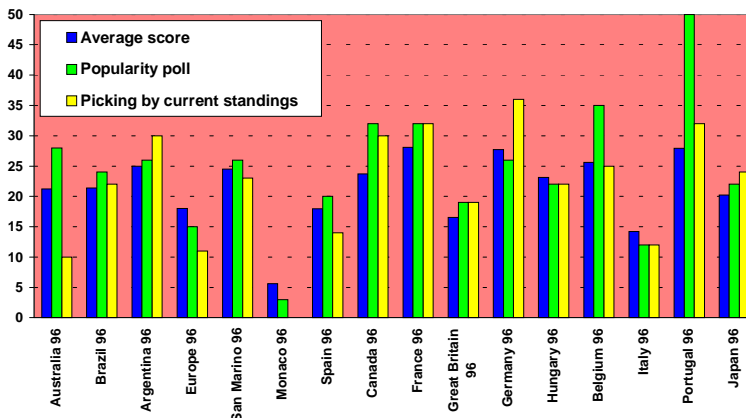
Average Pick 6 Score



Improvement...

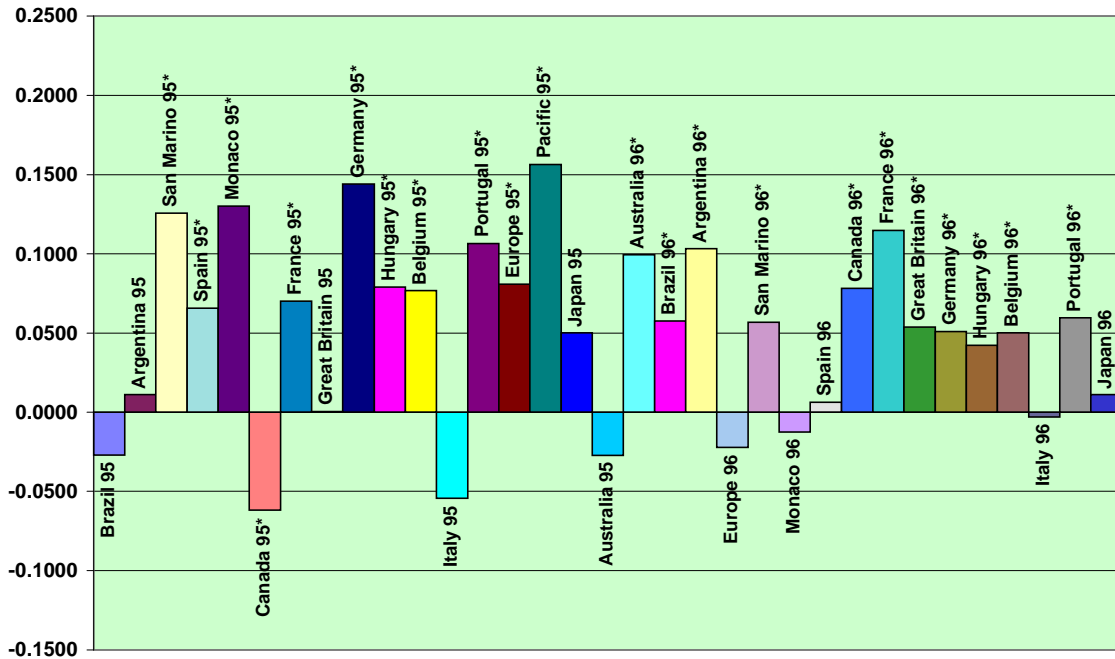
The chart above shows that average scores for participants have shown no pattern over three seasons. And the popularity poll (based on the most popular choices for each of the first six finishers) has been neither consistently better nor consistently worse than the average score. Yet, overall for three seasons, the popularity poll has taken 1033 points, while the average scores sum to 930.18. So perhaps the whole is smarter than the average of the parts.

1996 Pick 6 3 Ways to the Middle



The chart to the left shows that basing picks straight off the current driver standings (using the final 1995 standings for the first race) is less successful than the popularity poll (1996 totals: average 341, popularity poll 392, standings-based 342)

Correlation Between Score and Number of Previous Entries



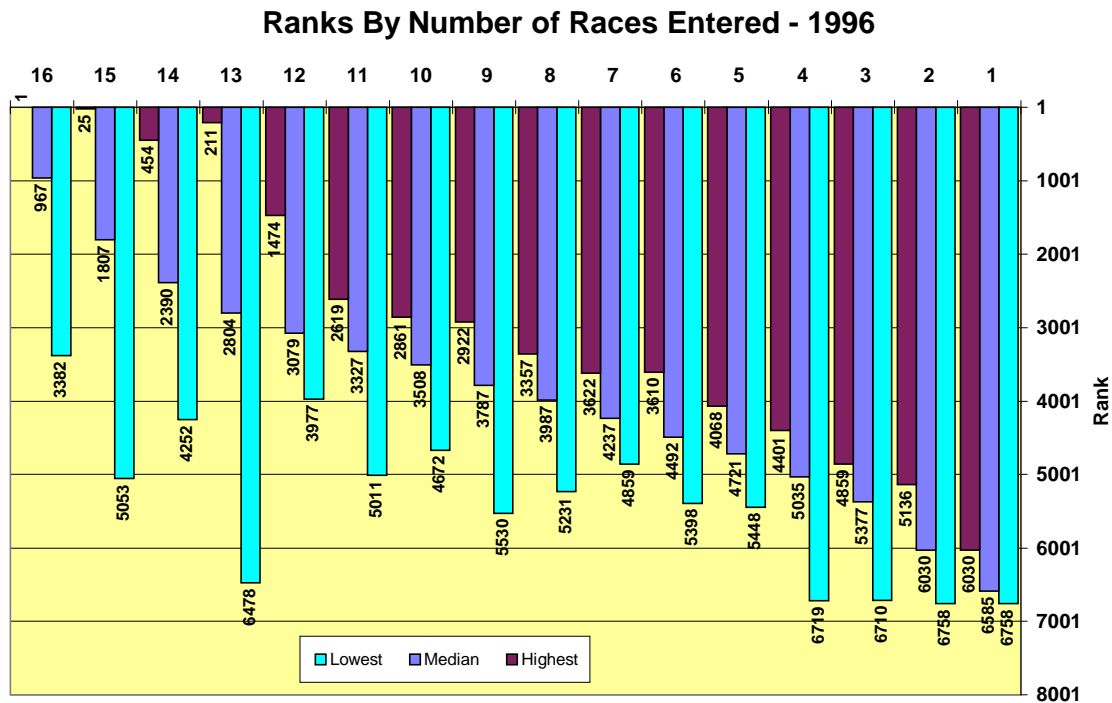
Experience...

Another way of checking for the value of experience is to compare the performance of SixPickers with more experience to those with less. The correlation coefficient is one way to put numbers to the relationship. The coefficients vary between +1, when knowing one value allows the exact prediction of the other, and as one goes up so does the other, through 0, when there is no relationship whatsoever, to -1, again allowing exact prediction, but as one goes up the other goes down.

The figure above shows the correlation between Pick Six score and number of previous entries since the beginning of the 1994 season. It covers the only 1995 and 1996 races, so as to allow a range in experience. Bars going up show experienced pickers generally did better than those with fewer entries; bars going down show they did worse. Asterisks show the races when the a correlation coefficient would be that large fewer than one time in 100 as a result of chance variation; in those cases we can be fairly sure there is a real relationship between experience and score. Only **Canada 95** among the negative relationships shows that level of statistical significance, while 21 other races show significant positive correlations. Clearly, the experienced pickers typically outscore the beginners. But a median correlation of +0.057 is hardly a powerful relationship; there's a lot more to picking winners than practice.

Winning...

What makes a winner in F1Pick6? More than any other factor, your total score for the year depends on how many races you entered. Miss one race, and your hopes of the championship are gone. The figure below (for the 1996 season) makes this relationship clear. Ranks for those entering all 16 races ranged from a high of 1st to a low of 3382nd, with a midpoint (median) of 967th. The highest rank for anyone entering only 15 races was 25th place, a considerable accomplishment; for 14 races it was 454th place. The median scores are strictly in order of number of races entered. The same order nearly holds for the best ranks for the number of races entered. Worst ranks, on the other hand, can be achieved with any number of entries.



The relationship has held for the last three years. The correlation coefficient showing the strength of the relationship between number of races entered and total score for the season was 0.978 in 1994, 0.980 in 1995 and 0.977 in 1996, out of a maximum possible value of 1.000.

Alternatives...

One way the scores could be calculated that would remove the advantage of entering more often would be to convert the usual scores to *z-scores*. This transformation also has the feature of adjusting scores for the difficulty of calling a particular race. The *z-score* transformation involves subtracting the average score from each score, and then dividing by the *standard deviation* of the distribution. Thus, if the average score was 24 points, anyone scoring exactly 24

points would have a z-score of 0 for that race, as would anyone who didn't enter. The standard deviation of the distribution (defined as the square root of the average of the squares of the differences of each score from the average score) will vary with the range of scores. Thus the z-score has the additional effect of taking into account how much better (or worse) you did than others. For example, the average score for Monaco 96 was 5.62, with a standard deviation of 4.54, and a maximum score of 30. This maximum would be transformed to a z-score of $(30 - 5.62)/4.54 = 5.37$. The average for France 96 was 28.09, with a standard deviation of 7.84, and a maximum of 60, which transforms to $(60 - 28.09)/7.84 = 4.07$. Scoring 30 on the Monaco 96 pick was a more outstanding performance compared to the other entries than scoring 60 on the France 96 pick, and is rewarded accordingly with a higher score.

The z-score transformation, (a *linear* transformation) has no effect on rankings in each race, and is only of interest in calculating total scores over a season (or more).

Ranking...

Compared to the correlation coefficients between number of entries and the conventional total score, which approached the maximum value of +1.00, the correlations between number of entries and total z-score were +0.107, +0.180, and +0.144 for 1994, 1995, and 1996, respectively. The overwhelming effect of number of entries is removed, and experience has about the same effect on total score as it was earlier shown to have on a typical race.

The table on the next page lists the thirty-five people with the highest total z-scores over the last three years of F1P6. Numero Uno is Chuck Abbott, by a large margin. Chuck has entered picks for every race in three years, and ranked among the leaders in 1994 and 1996 (5th and 3rd), and placed well in 1995 (345th). All of the top four have entered every time. The average number of entries is 37.7 of a possible 49; these are a practiced lot. The smallest number of entries was 15.

The ranks based on the standard scoring are quite unlike the combined z-total ranks; it seems that getting a high score consistently and beating the average performance consistently are somewhat different achievements.

These ranks are presented for your amusement only. Official bragging rights are still attached to the official ranks.

**The Top Thirty-five (1994-1996)
Ranked by Total z-Scores**

	Name	z Total	94 Rank	95 Rank	96 Rank	94 Entries	95 Entries	96 Entries
1	Chuck Abbott	26.08	5	345	3	16	17	16
2	David Kinloch	19.75	24	4	559	16	17	16
3	Sandy Englund	19.51	33	609	12	16	17	16
4	Jon Fairhurst	19.08	8	80	1088	16	17	16
5	Esa Vanhanen	19.04	NR	88	21	0	17	16
6	Antonio Jorge Marques	18.84	9	235	454	16	17	15
7	Christian Schakerl	18.46	24	NR	5	16	0	16
8	Greg Black	18.38	27	19	3547	16	17	9
9	Vinzenz Braune	17.53	150	192	49	16	17	16
10	Dick Carter	16.99	48	6	1213	16	17	15
11	Patrick D Taylor	16.87	NR	1788	211	0	8	13
12	Kimmo Venalainen	16.69	NR	21	61	0	17	16
13	Sean Casey	16.67	NR	60	72	0	17	16
14	Janne I Turunen	16.38	NR	NR	1	0	0	16
15	Eric Lessard	16.22	NR	235	10	0	17	16
16	Sam Hutchison	16.09	NR	319	125	0	15	16
17	Alvin Cheung	15.95	145	31	493	16	17	15
18	Damon Oriente	15.74	NR	72	46	0	17	16
19	Neale Type	15.70	12	NR	61	16	0	16
20	Bruce Fisher	15.04	77	97	277	16	17	16
21	Paul Smyth	15.02	385	1	864	14	17	16
22	Joe Hartman	14.94	225	187	179	14	17	16
23	R van Wijk	14.86	NR	108	79	0	17	16
24	Gary Klepper	14.85	21	178	433	16	17	16
25	Ken Chung	14.84	3	NR	NR	15	0	0
26	Leigh Haig	14.83	NR	80	255	0	17	16
27	Simon Rawlinson	14.81	NR	55	277	0	17	16
28	Mika Lilius	14.77	422	11	318	11	17	16
29	Dieter Kranzmueller	14.61	39	NR	67	16	0	16
30	Marko Ovaska	14.54	NR	NR	2	0	0	16
31	Kevin Kealy	14.46	NR	3	1510	0	16	14
32	Lorne Epp	14.33	193	319	58	15	17	16
33	David Richter	14.23	17	268	380	16	17	16
34	Phil Treide	14.20	NR	36	226	0	17	16
35	Steve Dodge	14.20	NR	21	203	0	17	16

Selecting...

One imagines entrants poring over race results, weather forecasts, track plans, and the output of the rumor mill (and perhaps a Tarot deck or a crystal ball) in order to get just the right prediction for that specific race. And yet if an entrant fails to file, the previous entry is carried over. If there is any science to the picks, one would expect the carry-overs to be less successful the second time than the first.

Using the z-score transformation outlined earlier to make direct comparison of races fair, the data show *absolutely no difference* on the average between the intended race score and the carry-over race score. If there are "horses for courses" their identification is not being successfully carried out. You might as well let your picks ride.

© David A. Wasserman, 1996

Permission to reproduce this document freely in unaltered form is granted.

Editor: David A. Wasserman
Writer: David A. Wasserman
Graphics: David A. Wasserman
Analyses: David A. Wasserman
Design, etc.: ...you get the picture

Appendix A courtesy of Paul Winalski

Software used to produce this document:

Microsoft Word for Windows 95
Microsoft Excel for Windows 95
SPSS 6.1.3 for Windows
SPF for Windows 1.0
Adobe Acrobat Pro 2.1
Netscape Navigator 2.0
Lview for Windows 3.1
Corel Draw 3.0

(Above are registered trademarks of their respective manufacturers.)

Data Mining and Data Analysis

David A. Wasserman, Ph. D.

Now the smaller business can get the type of information that large corporations and governments use to make their critical decisions. Twenty years of experience in creating information from data is available to you on a contract basis customized to suit your needs.

Data Mining

Computerized inventory and accounting systems, point-of-sale systems, customer records and other systems produce data that can be analyzed to pinpoint opportunities for efficiencies or expansion. With your expert knowledge of your business and my creative use of powerful analytic techniques, we can frame questions and find answers that will give your business an edge.

Surveys

Many small businesses try to track customer satisfaction through surveys, but lack the experience to put together effective survey designs or to analyze the responses. I can help you design questionnaires that I can analyze to give you reliable information presented in clear, concise reports. Your customers are trying to tell you something; I can help help you hear what they have to say.

Data Analysis

Are you assembling a team to perform in-house action research or program evaluations, but lack training in designing research or evaluation plans? I can assist in planning, data collection and analysis.

My Qualifications

I have been analyzing data since 1974. For the last 15 years, I have worked on the development, analysis and reporting of large-scale achievement testing programs in Alberta and across Canada. My work has also involved the design of numerous surveys, and the creation of effective designs for post-hoc analysis of existing data. I have also worked as a volunteer analyzing data for Edmonton Freenet, a not-for-profit community internet service provider. I have advanced degrees in Educational Psychology, and have worked as a door-to-door interviewer, an English teacher, a journalist, and a free-lance writer and book reviewer.

For more information, call me at (403) 465-7285 or send me electronic mail at davidwss@freenet.edmonton.ab.ca

Appendix A

For the sake of maintaining a self-contained document, Paul Winalski's F1 Pick6 **RULES AND NOTES** from the F1P6 web site follow:

• *Purpose of the Competition*

Formula One Pick Six (or, for short, F1P6) is a competition among the fans of F1 racing on the Internet to see who can do the best job predicting the points paying positions (1-6) of the Grands Prix. There's no money or prizes or anything, other than bragging rights.

• *How to participate*

Before each Formula One Grand Prix, fill in the form with your predictions (picks) for 1st through 6th place. If for some reason the form is not working for you, you can also send in your picks using electronic mail. Please use the form if at all possible.

• *Deadlines for picks*

Picks are due by the start of first untimed practice for the race. Usually this is 9:00 AM Friday morning, local time of the race site, but some races (e.g., Monaco) hold first practice on Thursday. We choose this time so that those who have access to early practice and qualifying information won't have an unfair advantage over those who don't.

If you use the Internet form, your pick is registered immediately.

If you send your pick by e-mail, it is your responsibility to see that it arrives in the pick6 mailbox by the due date. The processing programs go by the time the e-mail arrives in the mailbox, NOT the time that you sent the message. Picks arriving after the deadline will be flagged as late in the standings reports and are subject to review and possible rejection at the end of the season.

You may send in picks as far in advance of a race as you wish, and you may change your pick as many times as you wish, up to the dead- line. The last pick received and allowed will be the one that is used.

It is each player's responsibility to get their picks in on time. I send out a reminder message a week in advance of the due date, but because of the unreliability of network mail, you shouldn't depend on the reminders as your only way to remember to send in a pick.

• Picks may be carried over for one race. If you fail to send in a pick for a race, or if your pick arrives too late to be counted, your pick for the last race will be used again, but it will only be carried over once (e.g., if you fail to enter a pick for

race #3, race #2's pick will be used, but if you fail to enter again, race #2's pick will not be used for race #4).

• **Scoring**

You score points for each race based on how well your picks match the actual race results. For each driver that you picked, you score points based on where the driver finished versus where you predicted he'd finish:

didn't finish in the top 6	0 points
finished exactly where predicted	10 points
finished 1 place off from where predicted	6 points
finished 2 places off from where predicted	4 points
finished 3 places off from where predicted	3 points
finished 4 places off from where predicted	2 points
finished 5 places off from where predicted	1 point

For example, if you picked Schumacher to come in 3rd and he did finish 3rd, you'd score 10 points, but if he finished 5th, you'd score 4 points. If he finished 7th, you'd score 0 points.

Your total score for a race is the sum of the points scored for each driver in your pick.

If your pick includes the name of a driver who did not participate in the race, you do not score any points for that position. If a team substitutes one driver for another, you do not score anything if the substitute finishes in the points and you had picked the original driver.

If you accidentally pick the same driver more than once in your pick, you will not score any points for any of the positions in which that driver's name appears.

• **Determining the Pick6 Champion**

The F1 Pick6 Champion will be the player who scores the most total points over the season. All 16 Grands Prix will count towards the total.

If multiple players share the top score, they will be declared co-champions.

- ***Questions, Comments, etc.***

Questions, comments, and whatnot concerning the running of the F1 Pick6 competition should be directed to the F1 Pick6 Coordinator, Paul Winalski, who can be reached by e-mail at winalski@zko.dec.com. Do not use the pick6 mailbox for this purpose. The pick-processing program only understands picks and won't answer your questions.

If your e-mail address changes, please send mail to winalski@zko.dec.com informing the Coordinator of the change so he can update the distribution list.

If you decide to leave the competition, please send the Coordinator a note about it so he can remove you from the distribution list. He doesn't like sending out unwanted e-mail any more than you enjoy receiving it!

- ***More about late picks***

Ideally, e-mail picks that arrive late in the mailbox would be examined by hand and the time they were posted would be used to determine whether or not they really were sent in late. But with thousands of players participating, I do not have the time to do this. Therefore, the programs go by the time the mail message arrives in the pick6 mailbox. Picks arriving after this time will be provisionally accepted, but they will be flagged with an asterisk in the scoring reports. At the end of the season, any late picks that might have a bearing on the top positions in the competition will be reviewed and will be rejected if, in my judgement, they were in fact submitted late.

- ***Sending picks in advance.***

If you're going to be away from the Net, you can send in your picks for a race early. You don't have to wait until the race in question is the next Grand Prix.

You can change your pick for a race as often as you wish, up to the deadline for that race. I use the pick that was sent the latest, but before the deadline.

- ***Acknowledgments of picks.***

For your convenience and peace of mind, I send out an acknowledgment when I've processed your pick. For picks submitted ahead of time, I batch them up and process them just after I send out the Call for Picks for that race, 1 week before the deadline for picks. If at any time you want to check on the status of one of your picks, just send me email at winalski@zko.dec.com. Do not use the pick6 mailbox for questions about picks.

- ***Scoring and standings reports.***

After each race, the scoring program will send you a personalized status message showing how your pick scored for that race, your ranking for that race, your overall points total, and your overall rank.

An overall standings summary, showing the complete per-race and total points scoring for all players, is available in the Motorsports News International archive. It is also posted to newsgroup rec.autos.sport.info after each race. If you wish to be mailed a private copy, send me e-mail and I'll put you on a distribution list for it.

If you wish to see the detailed calculation of the Popularity Poll, send me e-mail and I'll put you on a distribution list to receive it.

All picks submitted by all players are public. If you wish to see the picks for any player and how they scored, send me e-mail.

- ***The driver Popularity Poll.***

You'll notice one "player" in the standings summary called Popularity Poll. The Popularity Poll's pick for each race is determined by treating each player's pick for the race as though it were the real race results, assigning F1 championship points to the drivers accordingly, and then taking the top 6 drivers as the pick. The Popularity Poll thus indicates the collective wisdom of all the players. In the last couple of years, it's performed well above average (a well-known psychological phenomenon known as the Delphi Effect).

Paul Winalski, winalski@zko.dec.com