PA 2013

HOW TO HANDLE EXPECTATION BIAS IN PRESENTIMENT EXPERIMENTS

A Recommendation

Jan Dalkvist[1], Julia Mossbridge[2] and Joakim Westerlund[3]

Summary

Here we reconsider expectation bias, with focus on how to handle it in experiments attempting to demonstrate presentiment. Presentiment is usually demonstrated by showing that significant physiological differences *precede* stimuli of different arousal levels, with all stimuli being presented in a randomized order with replacement. Often the direction of these differences suggests that physiological arousal is more likely to precede arousing rather than neutral stimuli. The possibility exists that such reactions can be explained as resulting from expectation bias. Expectation bias is based on the (false) notion that the likelihood of an arousing stimulus being presented grows as the number of consecutive calm stimuli increases (the gambler′s fallacy). Different ways of minimizing or avoiding the bias are discussed. On the basis of this discussion, our recommendation is to use analysis of variance (ANOVA) to separate the effect of the bias from the hypothetical presentiment effect, preferably at the trial-by-trial level. We also recommend ANOVA to be applied to each participant separately to avoid mixing within- and between- participant pre-stimulus effects, and to use a "counting" method to test for possible presentiment effects at the group level. The favoured method is illustrated using both a simulated one-participant example and real, multi-participant data. Finally, we anticipate that ANOVA can be performed to handle not only the expectation bias, but also other similar biases, like the so-called "hot hand" bias, in presentiment experiments as well as in conscious precognition experiments involving feed-back.

Keywords: presentiment, expectation bias, ANOVA, analysis of variance, the gambler′s fallacy, precognition

Imagine that you are running an experiment at the casino with a woman named Mary gambling on the roulette. Mary, like most people, is a victim of the "gambler's fallacy" and believes that, for each time the ball has dropped into a *red* slot, the chance increases that the ball will drop into a *black* slot the next time. Imagine also that Mary always bets on black, that she places a certain number of bets, say, ten in a row, and that for each successive time the ball has dropped into a red slot, her heart-beat will increase by one beat per minute until the ball drops into a black slot, whereupon her heart-rate will return to baseline. After the experiment, you compute the mean number of heart-beats at the moment before the ball dropped into a red slot and the mean number of heart-beats at the moment before the ball dropped into a black slot.

What is the statistically expected value of this difference? If you ask a group of statistically naïve people who, like Mary, are not familiar with the Gambler′s fallacy, they will certainly

---

[1] Department of Psychology, Stockholm University, Sweden (emeritus)
[2] Department of Psychology, Nothwestern University, Evanston, IL, USA
[3] Department of Psychology, Stockholm University, Sweden

tell you that the mean heart-rate will be higher before the ball drops into a black slot than before it drops into a red one, because – they argue – Mary´s heart-rate will reach a peak at that moment. By contrast, a group of statistically more sophisticated people, who *are* familiar with the gambler´s fallacy, will certainly tell you that the expected difference is zero. Their argument would be that the gambler´s fallacy is indeed a fallacy and that the mean number of heart-beats before the ball falls into a black slot could in the long run by no means be different from the mean number of heart-beats before the ball falls into a red slot, the stated reason being that each individual outcome is independent of all other outcomes and that two different outcomes have exactly the same likelihood to occur.

Which of the two groups is right? The surprising answer is that none of them and both of them are right. In short, assuming no unexplained phenomena are occurring, the statistically more sophisticated respondents are right in that the means of the heart-rates before the two alternative outcomes will be exactly the same in the long run. However, the statistically naïve respondents are also right: in the short run, when the length of a run is finite, the mean heart-rate will be higher just before a ball falls into a black hole than just before it falls into a red one. However, the reason suggested by the naïve respondents is wrong: the reason is not the fact that the arousal level tends to reach a peak just before the ball drops into a black hole, but is, in fact, much more difficult to understand.

We will now leave the casino for the moment and turn to the parapsychological laboratory. Here we will be acquainted with a certain kind of precognition experiment, akin to the above casino experiment, that was originally invented by Dean Radin and has come to be known as "presentiment experiments" (Radin, 1997). In such an experiment, instead of red and black slots, there are two types of pictures being shown to the participants, activating and calm pictures. The pictures shown are selected randomly, with replacement, from a pool. The participants´ physiological arousal at the moment before a picture is shown is measured. Sometimes heart-rate has been used as the measure of arousal, but all sorts of physiological reactions indicative of emotional arousal have been tested – including electrodermal activity, blood volume and pupil dilation. The hypothesis tested is that the arousal level becomes enhanced not only in response to the presentation of an activating picture, but also 2-15 seconds *before* the activating picture is presented, thus exhibiting a sort of physiological "presentiment" of the imminent presentation of an emotional picture that is different from the physiological activity preceding the presentation of a neutral picture. Although this hypothesis has received some support (for review, see Mossbridge, Tressoldi and Utts, 2012), several alternative interpretations have been discussed (Dalkvist and Westerlund , 2006; Mossbridge, Tressoldi, and Utts, 2012; Radin, 2004), including effects of the gambler´s fallacy, currently discussed here.

There is no doubt that in the long run, that is, in an infinitely long sequence of randomly ordered activating and calm stimuli, the average arousal level preceding calm stimuli will be equal to the average arousal level preceding activating stimuli – provided, of course, that no precognition is at work. Remarkably, however, this equality doesn´t hold for a *finite* sequence of calm and activating pictures, as suggested above. For such sequences, assuming the presentiment experiment is performed by a person who is subject to the gambler´s fallacy, the expected mean arousal level preceding activating stimuli are greater than the expected mean arousal level preceding calm stimuli. As would be expected, though, the difference between these means decreases as the length of the sequence increases, as demonstrated and explained in Dalkvist and Westerlund (2006).

Because the expected difference between the two mean arousal levels being compared decreases as the length of the sequence increases, the strength of the apparent bias will be dependent on the sequence length, short sequences tending to give rise to larger biases than long ones. Fortunately, however, the bias can be controlled: it can be reduced, altogether avoided, or its effects can be eliminated by statistical means. One of two purposes of the present paper is to compare these alternative courses of action and argue for choosing one of them: using statistical means to eliminate the bias. The other purpose is to demonstrate concretely, using simulated as well as real data, how this statistical correction method can be applied in practice.

That the bias is real and not merely a myth is a recent insight, dating back to the PA conference in Paris in 2002. At that conference decisive proofs were presented by our colleague Dick Bierman and two of us (JD and JW), using foremost simulated "paper-and pencil" demonstrations (Dalkvist & Westerlund, 2006); and, independently, by Jiri Wackerman (Wackerman, 2006), exploiting analytical mathematical methods rather than concrete simulated demonstrations. Before these proofs were advanced, most parapsychological researchers did not even contemplate the possibility that a bias of the present type might exist, or, if some knew of its existence, they thought it was illusionary or lacking in practical consequences.

Different ways of handling the bias

Our simulated examinations of the present bias – hereafter referred to as "expectation bias" – did not only reveal that the bias is real, but they indicated when, how and (at least partly) why this bias appears. For example, these examinations revealed that the bias can only be avoided altogether at a certain cost in terms of increasing sampling errors or increasing number of trials or participants, as will be seen below. Our simulations also revealed that different ways of designing the experiment or analyzing the data resulted in biases of different sizes.

To demonstrate some alternative ways of handling the bias, let us consider a very short presentiment experiment, involving only two pictures, one calm and one activating. This gives four possible sequences of pictures:

First one calm, second one calm.
First one calm, second one activating.
First one activating, second one calm.
First one activating, second one activating.

Now, let us assume that for each sequence, we have just one participant and that each participant believes in the gambler's fallacy and behaves accordingly. This situation is depicted in the left-hand column of Table 1. The participant who is to watch two calm pictures starts (like everyone else) with an arousal level of zero. After the first calm picture, he or she believes that the chance increases that the next picture will be activating, and his or her arousal level rises by one unit, meaning that, as far as this participant is concerned, the second stimulus is preceded by one arousal unit.

Table 1. Between-subjects analysis of expectation/arousal effects for activating and calm stimuli.

Arousal-modelled sequences	Stimuli: Activating $Sum(^1A)$	n_A	$Mean(^1A)$	Calm $Sum(^1C)$	n_C	$Mean(^1C)$
$^0C^1C$	0	0	-	1	2	1/2
$^0C^1A$	1	1	1	0	1	0
$^0A^0C$	0	1	0	0	1	0
$^0A^0A$	0	2	0	0	0	-
Sum	1	4	1	1	4	1/2
Mean	0.25	1	1/3 ≈ .33	0.25	1	(1/2)/3 ≈ .17
Bias = 1/3 - (1/2)/3 = 1/6 ≈ .17						

$^0A/^0C$ = zero arousal preceding an activating or a calm stimulus; $^1A/^1C$ = one arousal unit preceding an activating or a calm stimulus; $Sum(^1A)$ = sum of arousal units preceding activating stimuli; n_A = number of activating stimuli; $Mean(^1A)$ = mean of arousal units preceding activating stimuli; $Sum(^1C)$ = sum of arousal units preceding calm stimuli; n_C = number of calm stimuli; $Mean(^1C)$ = mean of arousal units preceding calm stimuli.

When we calculate the mean arousal level before activating pictures and the mean arousal level before calm pictures for each single participant and average each of the two resulting sets of means, we find that the mean of the individual average arousal levels before activating pictures (≈ 0.33) is larger than the mean of the average individual arousal levels before calm pictures (≈ 0.17). Thus, by calculating the means of the arousal levels before activating and calm pictures, we get a bias, which in this particular case is substantial. This bias would, of course, still exist in between-participant comparisons, assuming most participants held the same bias.

Table 2. Within-subjects analysis of expectation/arousal effects for activating and calm stimuli.

Arousal-modelled sequences	Stimuli: Activating $Sum(^1A)$	n_A	$Mean(^1A)$	Calm $Sum(^1C)$	n_C	$Mean(^1C)$	$Mean(^1A)$ - $Mean(^1C)$
$^0C^1A$	1	1	1	0	1	0	1
$^0A^0C$	0	1	0	0	1	0	0
Sum							1
Mean							1/2 (Bias)

$^0A/^0C$ = zero arousal preceding an activating or a calm stimulus; ^1A = one arousal unit preceding an activating stimulus; $Sum(^1A)$ = sum of arousal units preceding activating stimuli; n_A = number of activating stimuli; $Mean(^1A)$ = mean of arousal units preceding activating stimuli; $Sum(^1C)$ = sum of arousal units preceding calm stimuli; n_C = number of calm stimuli; $Mean(^1C)$ = mean of arousal units preceding calm stimuli.

A within-participant design, in which the means of an individual's arousal levels before calm and activating stimuli are compared with one another, is not a better method to handle the bias. This is because even though sampling errors would become smaller, expectation bias doesn´t get any smaller in a within-participant analysis compared to a between-participant analysis. This is shown in Table 2, where data from Table 1 is subjected to a within-participant analysis. In this analysis, the CC sequence and the AA sequence from Table 1 (the first and the last sequence) have been dropped, as can be seen. The reason is, of course, that in a within-participant analysis, not only undefined parameters (ratios with zeros in the denominator) have to be excluded, but also the whole sequence consisting solely of C stimuli and the whole sequence consisting solely of A stimuli, so that each sequence will contain at least one A stimulus and at least one C stimulus. As can be seen, the within-participant analysis leads to an even larger bias than the between-participants analysis does[4].

To shrink the size of the bias, there is, a better – in fact a much better – method of computing the overall mean difference between arousal levels before activating and calm stimuli: to pool all data across participants before averaging. This method is illustrated in Table 3. When calculating the mean arousal level before activating pictures and before calm pictures, we do not care about which participant the data come from. Thus, we just sum all arousal levels before activating pictures (= 1) and divide the sum by the total number of activating pictures (= 4). The corresponding mean arousal level is one-fourth. We do the same computations for the arousal levels before calm pictures. In the present example, we get exactly the same mean as that obtained for the activating pictures. Thus, the difference between the two means is zero, and we do not get any bias at all.

Table 3. Analysis of expectation/arousal effects for activating and calm stimuli using complete and balanced pooling.

Arousal-modelled sequences	Stimuli: Activating $Sum(^1A)$	n_A	Calm $Sum(^1C)$	n_C
$^0C^1C$	0	0	1	2
$^0C^1A$	1	1	0	1
$^0A^0C$	0	1	0	1
$^0A^0A$	0	2	0	0
Sum	1	4	1	4
Mean	1/4		1/4	
Bias = 1/4 -1/4 = 0				

$^0A/^0C$ = zero arousal preceding an activating or calm stimulus; $^1A/^1C$ = one arousal unit preceding an activating or a calm stimulus; $Sum(^1A)$ = sum of arousal units preceding activating stimuli; n_A = number of activating stimuli; $Sum(^1C)$ = sum of arousal units preceding calm stimuli; n_C = number of calm stimuli.

Unfortunately, this complete absence of any bias is exceptional. In order to avoid the bias completely doing the present calculations, it is necessary that (a) all possible sequences occur in the experiment and (b) each sequence occurs the same number of times. For practical reasons, however, such complete, balanced pooling can generally not be achieved. As can be seen from Table 4, if there is an equal probability for calm and activating pictures, we need at least 4 participants for a sequence length of 2 trials, 8 participants for a sequence length of 3 trials, 16 participants for a sequence length of 4 trials, and so on. Soon, however, the number of participants will be impractically large, as can be seen at the bottom lines in Table 4.

[4] However, when the data are exactly the same in the two methods, the values of the expectation bias will also be exactly the same, even though expectation bias can be more easily confirmed statistically using the within-participant analysis, due to the sampling errors being smaller using this method.

Table 4. Minimum number of participants required for the bias to disappear when all data are pooled across participants.

Sequence length (Number of trials for each participant)	Number of participants required
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
20	$1.049 * 10^6$
30	$1.074 * 10^9$
40	$1.100 * 10^{11}$

There are, however, also other ways of avoiding the bias. One is to refrain from calculating means and just calculate and compare the associated sums. That the sums become equal in spite of the fact that the corresponding means differ can be seen from Table 1, where both of the two arousal sums are equal to unity. Still another way of avoiding the bias is simply to let each participant be presented with only one stimulus.

In Table 5, four different ways of handling the bias are evaluated in terms of drawbacks and lack thereof. As indicated in this table, the common drawback of the three above-mentioned methods of avoiding the bias altogether is their requirement of many participants or individual trials. Using sums instead of means require a large number of participants or individual trials to make up for random errors resulting from unequally large samples of activating and calm stimuli, and pooling data to form a complete and balanced design requires a large number of participants to fill the large number of different sequences, as we saw in Table 4. Similarly, using different participants in different single-trial experiments requires a large number of participants to keep the inter-individual variation at a sufficiently low level to uncover a possible presentiment effect.

To get a small bias, pooling data but permitting the design to be incomplete and unbalanced is a much better strategy than first calculating means on the individual level and then calculating group means on the basis of these individual means, as the bias becomes much smaller in the pooled design than in a design based on individual means. Nevertheless, this method is not quite satisfactory, for two reasons. Firstly, we don´t know exactly how big the bias is, which, in turn, makes it impossible to perform an exact statistical test of the presentiment effect. Secondly, when using conventional statistical methods to test for any presentiment effect, one runs the risk of violating the assumption of independent observations, as the within-participant responses are statistically dependent, even though the between-participant responses are not.

Uncertainty may also affect the third of the methods listed in Table 5: to test if there really is any bias, using, for example, regression analysis. Thus, even if there isn't any statistically significant bias, but only a tendency for a bias to occur, one doesn't know if, and how much, the bias will affect a subsequent statistical test of the hypothetical presentiment effect.

The fourth and final method has no obvious drawback to it, and is therefore in general superior to the three other ones. The goal of this method is two-fold: (1) to test and eliminate any possible expectation effect and (2) to test for the existence of a presentiment effect, independently of any expectation bias, using analysis of variance (ANOVA).

Table 5. Four Ways of Handling the Bias

Goal	Method	Drawback
Avoid the bias	1. Use sums instead of means 2. Pool data in a complete and balanced design 3. Use different participants in different trials	Resource-consuming: requires many participants or individual trials
Reduce the bias	Pool data in an incomplete and unbalanced design	Uncertainty; may violate the assumption of independent responses
Find out if there is a bias	Test if there is any statistically significant bias	Uncertainty
Eliminate the effect of a possible bias	Test and eliminate the effect of any bias using a two-way analysis of variance	No obvious

A simulated ANOVA

This last method will now be described in more detail using a simple simulated example. In the above simulated illustrations, the number of trials has been restricted to two. In the simulation to be presented below, however, the number of trials has had to be larger. This makes it necessary to specify more exactly how the arousal level varies as a function of successive calm and activating stimuli.

In the introductory casino experiment, Mary's heart-rate was governed by a linear arousal model: her arousal level increased linearly as a function of the number of successive drops of the ball into a red slot. An alternative growth model would be one in which arousal grows as a positively accelerated function of the number of previous calm pictures in succession, such as an exponential function. However, the most realistic growth model should probably have a sigmoid form, meaning that arousal grows only slowly or not at all at the beginning of a series of calm stimuli and increasingly fast as the number of calm stimuli increases up to some inflexion point at which the curve levels off. A simplified version of the sigmoid arousal model is a binary model. In that model, instead of increasing monotonically in a series of calm pictures, as in the above models, arousal increases from base-line with a certain amount at the first picture in a series of calm pictures and remains at that level until an activating picture resets the level to base-line.

However, for testing the expectation bias using analysis of variance (or some similar statistical method such as regression analysis) any model according to which the arousal level increases in response to some or all calm stimuli and drops in response to an activating stimulus can be assumed. This is true even when different participants have different expectation functions.

In demonstrating how presentiment data can be handled using an ordinary independent measures ANOVA we have constructed a simple example involving only one participant and 12 activating (A) or calm (C) pictures:

C,C,C,A,C,A,A,C,C,C,C,A.

For simplicity, we have chosen the linear arousal growth model rather than any non-linear such model to describe the arousal level as a function of successive calm and activating stimuli. That is, for each calm stimulus that the participant encounters, we add a new arousal unit until an activating stimulus appears, whereupon the arousal level drops to base-line, which in our example is set to zero. Furthermore, we have assumed that the participant has an infallibly precognitive ability: his/her arousal level increases by one unit whenever the next picture is activating.

As can be seen from Table 6, we have entered three sources of arousal: the preceding stimulus, the next stimulus and an error generator, giving rise to random fluctuations. If there is an expectation bias, a correlation exists between the type of preceding stimulus and arousal: a calm stimulus, coded as"0", has the effect of increasing the arousal level by one unit, whereas an activating stimulus, coded as "1", resets the arousal level to baseline ("0"). And if there is a presentiment effect, a correlation exists between the next stimulus and the arousal level: an activating, but not a calm, stimulus has the effect of increasing the arousal level by one unit. Finally, we have the error generator, which in our model, for simplicity, has been assumed to generate the same absolute value throughout: 0.25, which randomly shifts between plus and minus.

Table 6. Arousal Generation in a Simulated Presentiment Experiment

Preceding stimulus (position number)	Source of arousal			Next stimulus (position number)	Preceding stimulus (position number)	Source of arousal			Next stimulus (position number)
	Prec. stim.	Next stim.	Error generator			Prec. stim.	Next Stim.	Error generator	
C(1)	1	0	-.25	C(2)	C(2)	1	0	.25	C(3)
C(3)	1	1	-.25	A(4)	A(4)	0	0	.25	C(5)
C(5)	1	1	.25	A(6)	A(6)	0	1	-.25	A(7)
A(7)	0	0	.25	C(8)	C(8)	1	0	.25	C(9)
C(9)	1	0	.25	C(10)	C(10)	1	0	.25	C(11)
C(11)	1	1	.25	A(12)					

Table 7 summarizes our simulated data in the form of three variables to be used in the ANOVA, type of succeeding and preceding picture, respectively, being independent variables and arousal, including error, the dependent variable.

Table 7. Variables used in the Single Person ANOVA

Trial no.	Preceding picture	Succeeding picture	Arousal
1	0	0	.75
2	0	0	1.25
3	0	1	1.75
4	1	0	.25
5	0	1	2.25
6	1	1	.75
7	1	0	.25
8	0	0	1.25
9	0	0	1.25
10	0	0	1.25
11	0	1	2.25

Table 8 and Figure 1 show the results of applying an independent two-way ANOVA to the variables shown in Table 8. The significant effect of the type of preceding picture reveals an expectation effect. There is also a significant presentiment effect, even though this effect is weaker than the expectation effect. There is no significant interaction effect, but a tendency for the succeeding picture effect to be larger when the preceding picture is calm than when it is activating. Thus, our analysis answers two questions simultaneously. One is if there is any expectation effect, and the other if there is any presentiment effect, independent of any expectation effect. The answer to both questions is clearly affirmative.

Table 8. Results From an Independent Two-Way ANOVA of a
Simulated Presentiment Experiment

Picture	F	df	p
Preceding	46.83	1/7	.0002
Succeeding	19.29	1/7	.003
Interaction	1.76	1/7	.226

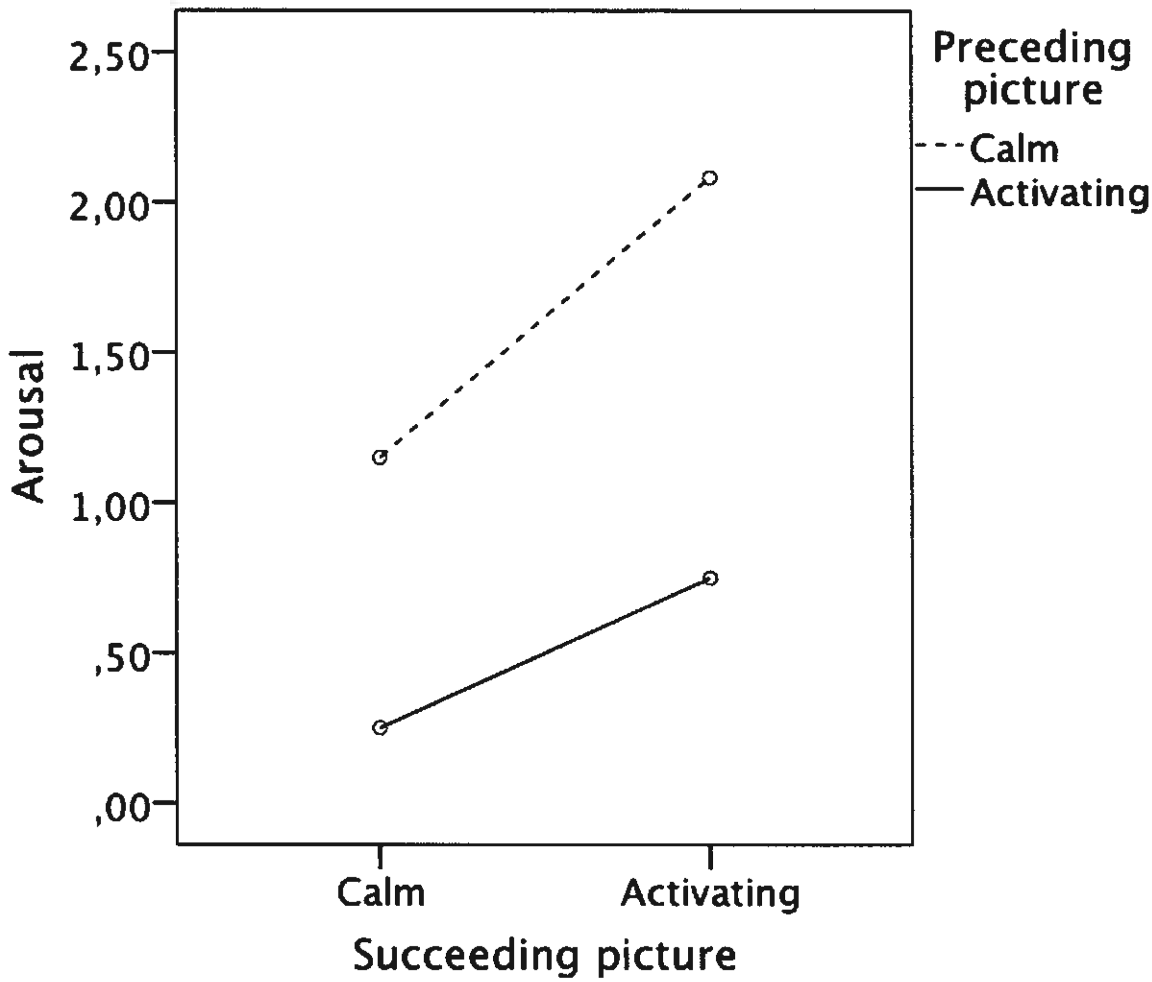

Figure 1. Arousal plotted against the type of succeeding stimulus for each type of preceding stimulus in the simulated presentiment model. "1" indicates an activating picture and "0" a calm picture.

There are several different ways of applying ANOVA to presentiment data when there is more than one participant. In such cases, one has to choose between a pooled analyses or a within-participant analysis, as defined previously, where data are analyzed collectively, and a between-participant analysis based on individual ANOVAs. Although the pooled analysis can be expected to yield a small bias, it violates a fundamental assumption in ANOVA: that of independent observations, even though this weakness can be handled by simulation techniques. Another problem is associated with the within-participant method: we don´t know how well the individual average arousal levels before calm and activating stimuli represent the trial-by-trial data. More importantly, in a within-participant analysis we mix different individual response patterns in an unknown way.

In general, we think, the between-participant method based on individual ANOVAs is preferable to the two collective methods discussed above. If one follows the statistically conservative approach and chooses to perform N individual ANOVAs on the trial-by-trial data (one ANOVA for each participant), the proportion of significant main effects of presentiment can be used to determine the statistical significance of presentiment in the given set of participants. The only flaw with this approach is that the result of this significance testing cannot be used to infer presentiment in the general population, only the particular participants who were tested.

Here we perform sample ANOVA analyses on a subset of data taken from one of our (JM's) guessing tasks. Briefly, participants were presented with four neutral images, and were given as much time as they liked to select the image that they guessed would be presented to them later. Once they made their selection, the software randomly chose one of the four images and presented that image to the participant. Trials were divided according to correctness. Post-stimulus responses to the visual feedback indicating that a guess was correct were generally in the direction of arousal (a positive change in skin conductance), while less arousing or neutral responses occurred after feedback indicating that a guess was incorrect. The set of data from which we took this subset were analyzed and reported elsewhere (Mossbridge, Grabowecky and Suzuki, 2011). – Note that the data shown here are not to be used in any meta-analyses of presentiment as they have been published elsewhere.

For purposes of illustration, we used the subset of the data that satisfied two criteria. One was the occurrence of the postulated expectation bias, that is, a negative correlation between the activation level of preceding stimuli and arousal. The other criterion was the occurrence of a presentiment effect, that is, a positive correlation between the activation level and arousal. Eighteen females (27%) and 11 males (24%) satisfied these criteria. However, due to gender differences already described elsewhere (Mossbridge, Grabowecky and Suzuki, 2011), we confined ourselves to consider only the female data here.

Table 9 shows the results from the eighteen females satisfying our two inclusion criteria. As can be seen, for most of the participants, the F-ratio for the preceding stimulus didn't reflect any expectation effect. However, for four of the participants (Participants 5, 11, 14 and 16) there was an apparent effect, even though only one of the F-ratios obtained (for Participant 11) was significant.

As can further be seen, the succeeding pictures exhibited a similar pattern: most of the F-ratios didn't show any presentiment effect at all, but some of them (for Participants 2, 4, 5, 9 and 10) did so, and two of the F-ratios (for Participants 4 and 10) reached significance.

If the participants had not been selected, but had made up the whole sample/population, it would have been possible to test whether these two participants made the presentiment effect to be significant at the group level, using a binominal test. Such a test shows, however, that four, rather than two, significant individual results at the 5% level would have been required to make the group results significant at the same significance level.

Participant 5 is of particular interest, as this participant clearly shows how the expectation effect and the presentiment effect can be separated from each other by using two-way ANOVA.

Table 9. Results from two-way ANOVAs applied to selected individual presentiment data from Mossbridge et al., 2011

Participant	Preceding pictures		Succeeding pictures		Interaction	
	F	p	F	p	F	p
1	.98	.33	.009	.92		
2	.39	.54	2.59	.13	1.76	.20
3	.00	.98	.97	.34	.02	.89
4	.30	.59	6.17	.09	2.83	.11
5	3.03	.10	2.54	.13	.64	.43
6	.07	.80	.80	.38	3.34	.08
7	.37	.55	.09	.77	.27	.61
8	.82	.38	.39	.54	.93	.35
9	.00	.95	1.82	.19	.12	.73
10	.21	.65	5.89	.02		
11	6.06	.02	.31	.58		
12	.31	.58	.40	.54	.02	.89
13	.22	.65	.39	.54	.39	.54
14	3.58	.07	.89	.77	.69	.42
15	.31	.58	.95	.34	.01	.93
16	1.71	.21	.00	.99		
17	.11	.75	.80	.38	.11	.74
18	.07	.79	.25	.62	.01	.94

However, the most important advantage of using a two-way ANOVA at the individual level is not any particular finding (apart from the number of significant presentiment effects) but the guarantee that the presentiment effects are not contaminated by expectation effects or any other similar effects. In other words, the two-way ANOVA serves as a filter for such effects, making the final, global statistical test reliable.

Discussion

A recent meta-analysis of 26 reports published between 1978 and 2010 revealed a highly significant, if small, presentiment effect (Mossbridge, Tressoldi & Utts, 2012). This effect persisted, and in fact became larger, when only studies finding no evidence for expectation bias were examined.

All three authors of the present paper have pointed out that presentiment data should be tested for expectation bias. The above meta-analysis indicated that about 50% of the studies satisfied this recommendation. Even if it is the case that expectation bias does not strongly influence presentiment results, it is critical to examine presentiment data for this bias and to report the results of these examinations. As we mentioned earlier, even non-significant expectation bias can result in a trend that influences the measured presentiment effect.

It may be argued that instead of applying ANOVA to individual participants, one should pool data across participants, because the bias will be minimal using this method. However, as pointed out previously in this paper, if using ANOVA, the basic assumption of independent measures will be violated in this case, although this problem can be overcome using simulation techniques. As we illustrate here, this problem can also, and more easily, be overcome by performing independent ANOVAs on each participant´s data, then determining the proportion of participants that have significant main effects for presentiment. If this proportion is higher than would be expected by chance, this result would suggest that for these participants, presentiment is occurring. For instance, if out of 20 participants more than three participants showed a significant main effect of presentiment, this would suggest that presentiment was occurring for those participants ($\alpha=0.05$), as indicated by a binominal test.

It is important to note that the single-trial method can be used to avoid all discussion about expectation bias. However, few scientists have access to the number of participants required for using this method. Moreover, we need a standard statistical method, such as ANOVA, to handle archieved data sets.

One drawback of the trial-by-trial ANOVA analysis we describe is that it only takes into account bias resulting from previous trial. However, it would be trivial to add more independent variables into the AOVA to account for trials 2 and 3 previous to the current trial, though this solution still does not consider potential differential effects of particular types of "runs." For instance, an oscillating sequence of trials (A,C,C,A,C,C,A,C) could conceivably produce an expectation for an upcoming calm stimulus (the next item in the repeated series). However, it seems likely to us that such effects will not dominate, as such repeating sequences are generally rare in well-randomized series.

Finally, we note that the gambler´s fallacy is not the only bias that potentially may create a bias that can influence presentiment measures. An opposing bias, the "hot-hand" bias, has been described elsewhere (e.g., Roney and Trick, 2009). The hot-hand bias is primarily held by people who believe that a sentient agent controls the outcomes of the events in a series, for instance, points in a basketball game. A participant who is subject to the hot-hand bias would expect a series of activating images to continue. However, it is not clear whether this bias would produce an increase in the type of expectation bias effect examined here (if arousal continues to increase before each activating image), or if it would produce an effect in the opposite direction (if arousal drops as each expected activating image is presented). Whether or not participants hold the hot-hand bias, the gambler´s fallacy bias, or neither, any bias that is present can be separated from an existing presentiment effect using the trial-by-trial ANOVA analysis described here. Further, this trial-by-trial ANOVA analysis could potentially be used to separate sequence effects from precognition effects in a standard conscious precognition experiment.

References

Dalkvist, J., Westerlund, J. & Bierman, D. (2002). A computational expectation bias as revealed by simulations of presentiment experiments. Proceedings of Presented papers: The 45th Annual Convention of the Parapsychological Association, 62-79.

Dalkvist, J., & Westerlund, J. (2006). A bias caused by inappropriate averaging in experiments with randomized stimuli. *Journal of Parapsychology, 70*, 233–254.

Mossbridge, J., Grabowecky, M. & Suzuki, S (2011). Physiological markers of future outcomes: Three experiments on subconscious psi perception during concurrent performance of a guessing task. Proceedings of the 54th Annual Convention of the Parapsychological Association:17.

Mossbridge, J., Tressoldi, P, & Utts, J. (2012). Predictive anticipatory activity preceding seemingly unpredictable stimuli: A meta-analysis. *Frontiers in Psychology,* 3:390.

Radin, D. (1997). Unconscious perception of future emotions: An experiment in presentiment. *Journal of Scientific Exploration, 11*, 163-180.

Radin, D. (2004). Electrodermal presentiments of future emotions. *Journal of Scientific Exploration, 18*, 253-273.

Roney, C. & Trick, L. (2009). Sympathetic magic and perceptions of randomness: The hot hand versus the gambler´s fallacy. Thinking and Reasoning, 15, 197-210.

Wackerman, J. (2002). On cumulative effects averaging artefacts in randomized S-R experimental designs. Proceedings of Presented papers: The Parapsychological Association 45th Annual Convention of the Parapsychological Association, 293-305.