

## Head-On-Pillow Defect – A Pain in the Neck or Head-On-Pillow BGA Solder Defect

Chris Oliphant, Bev Christian, Kishore Subba-Rao, Fintan Doyle, Laura Turbini,  
David Connell and Jack Q. L. Han  
Research In Motion, Waterloo, Ontario, Canada

The head on pillow defect is becoming more common. This paper describes one such occurrence for an OEM and explains how it was dealt with. In this particular case it was solved by application of problem solving skills by the OEM, component supplier and the solder paste provider.

### Introduction

Head on pillow (HOP), head in pillow, pillow defect, ball in socket, hidden pillow, foot in mud, ball in cup or the much preferred pain in the neck, when one is in the midst of an episode of it, is a difficult challenge to solve. In answer to the first “why?”, one can state that after reflow the re-solidified solder ball of the BGA or CSP is in physical contact with the flux residue, oxide coating or the actual solder of the re-solidified solder paste on the pad, but has not actually coalesced into a single solder ball with co-mingled solder. Usually there is at least some solder to solder direct contact and therefore the circuits containing the component in question pass electrical testing. In answer to the second “why?”, one might suggest: smaller solder deposits, higher processing temperatures, longer pre-liquidus thermal exposure, more tenacious oxide films on molten solder, smaller spheres on array components, package warpage<sup>1</sup>, solder ball dopants and/or reflow type (convection vs. vapor phase).<sup>2</sup>

The third “why?” expands out because of all the avenues suggested above. Many of them are affiliated with the fact that the HOP defect is seen more often with lead-free soldering, which requires higher temperatures and/or longer exposure to heat and has also lead to a resurgence of vapor phase soldering. The small solder deposits and small solder balls are a consequence of the drive to “smaller, cheaper, faster”. One could say this is an answer to a fourth “why?”. Jenson<sup>3</sup> has shown that the high surface area to volume ratio of small solder deposits can lead to a phenomenon called “graping” because of the exhaustion of the small amount of flux in miniscule solder paste deposits. The “fifth” why is the thickness of the solder paste deposit itself.

Why would components warp? Did it never happen before? There are two reasons for its increase. A plastic packaged integrated circuit is essentially a silicon wafer embedded in a plastic. Because of the demand for “smaller” components, manufacturers are back-grinding the silicon to make thinner components. This is changing components from having a relatively stiff backbone to something that resembles a bi-material strip – sort of like the bimetallic strip in your toaster. And of course when that is heated, the difference in coefficients of thermal expansion of the two materials leads to the metal strip bending over, completing a circuit, shutting off your toaster’s heating element and releasing your toast.

The dopants are an attempt to make solder joints that are more likely to stand up to temperature cycling and/or shock, the latter usually from drops of handheld devices. Karl Seelig of AIM has stated<sup>2</sup> that they have found that 30 ppm of magnesium, 400 ppm of nickel or cobalt, 500 ppm of indium or 1000 ppm of antimony “directly affect standard solder paste flux chemistries”. Since this was stated in an article on HOP, this implies that these dopants can have a significant effect on this particular defect.

So either because of inconsistently smaller solder paste deposits and/or smaller solder balls on the array component, usually in concert with component warpage, the solder of the solder ball and the solder in the solder paste melt as either the edges or the center of the component lift up out of the solder paste. The flux in the solder paste removes the oxide on the solder spheres of the solder paste and the oxide on the board pads or dissolves the OSP coating, if present. The flux is consumed and also evaporates/sublimes and only flux residue remains. Depending on the quality and quantity of the flux residue, the surface of the solidifying solder from the solder paste may or may not grow a new overarching oxide sheath. Certainly the lead-free solder ball, raised above the flux in temperatures of 235°C – 260°C oxidizes. As the circuit board starts to cool, the component starts to flatten out and a solidified or semi-solidified solder ball comes down on the solder on the pad and just kisses it or perhaps even deforms it if the solder is still not completely solidified. This explains the head on pillow analogy.

Several suggestions have been made to deal with the problem. One suggestion is to use an inert atmosphere – less oxidation or re-oxidation of the solder surfaces and therefore less consumption of the available flux. However, many sites do not have access to nitrogen or even have reflow ovens outfitted for such a retrofit.

Shea<sup>1</sup> also suggests a faster ramp rate, preserving flux. This is possible to one extent or another, but will it be enough for a particular situation? Certainly it would decrease the amount of time that a component would have to warp, but is the time or just the temperature the major factor? And what effect will this have on solder paste on pads connected to traces with high thermal conductivity?<sup>4</sup> A faster ramp rate should certainly decrease oxide formation and flux decomposition.

Another suggestion of Shea with similar limitations is to use a more thermally robust flux and or higher activity<sup>5</sup> in the solder paste. One solder paste manufacturer speaks of the higher heat resistance (thermal robustness) and also mentions quicker wetting reaction speeds<sup>6</sup>.

Lathrop<sup>7</sup> suggests determining the warpage potential of the virgin, unballled components and then printing solder paste on the under bump metallization with stencils having varying aperture sizes to form different sized balls on the array devices to take into account the expected warpage of the component.

And a final suggestion by Shea<sup>1</sup> that can almost always be applied, is to open the stencil apertures by one or 2 mils. This helps in three ways – more solder paste, a more favorable surface area:volume ratio and a more favorable aperture ratio for paste release.

Head on pillow defects centered in the middle of the component/pad array are most likely due to the component warping in a convex fashion, as seen from the top of the component. Head on pillow defects around the periphery of the component/pad array are most likely due to the component warping in a concave fashion, as seen from the top of the component. Random HOP defects may be due to one of more of:

- 1) thermal traces and vias
- 2) smaller paste deposits
- 3) smaller component solder balls

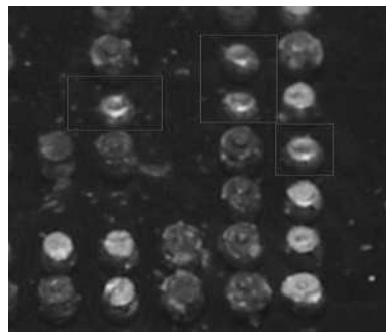
## Results

### Waking up with a HoP

HOP defects for BGA processors for a new product were detected in the initial builds. Depending on which BGA joint is impacted, HOP defects can escape electrical test. Pry-off and microscopic inspection is the most effective detection method, hence containment is prohibitively expensive.

In August 2007, pry-off and microscopic inspection of BGA solder joints for 100 prototype boards by the OEM lab team revealed 5 samples exhibiting random HOP defects (**Figure 1**), with no clustering patterns. Joints adjacent to HOP defects appeared normal. Cross sectioning was also used to confirm this. The upper bound estimate (95% confidence) of the true HOP defect rate was 10%.

The first order of business was to collect data. The circuit boards had a high temperature OSP finish. The solder paste was from a first tier supplier and was a SAC 305, Type 3 solder paste. The stencil was a 4.5 mil, E form stencil. The reflow profile was a straight ramp (SAC305, OEM standard oven recipe for all products). The BGA in question was 1 mm thick, 0.5mm pitch, with 300 micron diameter SAC105 solder balls.



**Figure 1 Evidence of solder balls on a reflowed component where the solder balls were not soldered to solder on the board pads.**

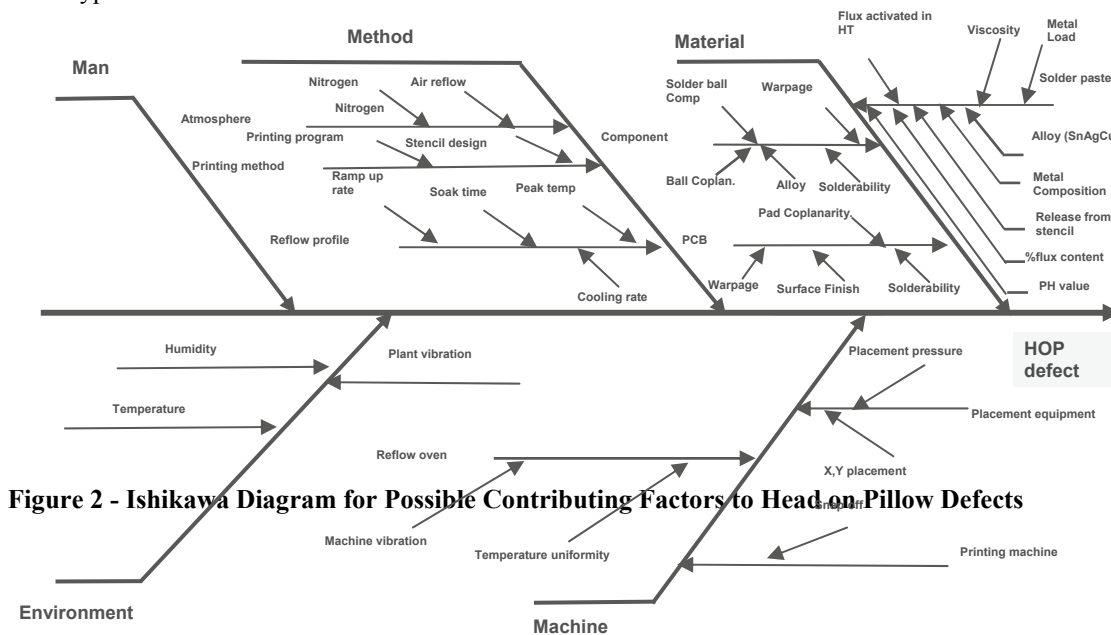
The first measurements taken were the warpage of the circuit board, by means of Thermoire™. The board in the area of the component exhibited either concave or complex shape. The co-planarity was measured at 164 +/- 25 microns. Two iterations of board design brought this down below 60 microns.

Many components were then measured for warpage and the results are shown in **Table 1**. At this point in time a warpage of less than 90 microns was considered acceptable. Any range containing 90 microns or any average at or above this amount are in italics.

**Table 1 - History of BGA Measured Warpage at Reflow Temperature**

Part	Lot #	Max Warpage Temp'	Range (Microns)	Average (Microns)	Std Dev
1	a	Var'	<i>80-100</i>	<i>91</i>	6
1	b	250	<i>73-90</i>	82	7
1	c	220	<i>73-90</i>	83	8
1	d	180	<i>47-90</i>	73	16
1	e	240	<i>61-107</i>	78	17
1	f	250	<i>65-76</i>	69	4
1	g	250	<i>63-82</i>	70	7
2	a	250	<i>96-106</i>	<i>101</i>	4
2	b	250	<i>101-114</i>	<i>107</i>	6
2	c	250	<i>85-91</i>	89	3
2	d	250	<i>79-92</i>	84	6

A team was formed to solve the problem. An Ishikawa diagram, covering a span of possible contributing factors, much bigger than the focused items mentioned in the introduction, was drawn up. See **Figure 2**. Based on failure analysis, cross sections, defect patterns and inputs from subject matter experts, 2 major failure mechanisms and contributing root causes were hypothesized.



**Figure 2 - Ishikawa Diagram for Possible Contributing Factors to Head on Pillow Defects**

**Hypothesis 1 = Surface oxidation prior to wetting**

- Flux in solder paste burns off prior to wetting resulting in surface oxidation of solder balls preventing formation of a proper metallurgical bond.
- The melting temperature difference between solder balls (SAC 105, 227°C m.p.) and solder paste (SAC 305, m.p. 217°C) results in solder paste remaining in liquidus phase longer than solder balls, resulting in oxidation prior to wetting.

- c) High gold content in intermetallics on component solder ball surface results in poor soldering and absence of metallurgical bond.
- d) Solder paste formulations differ in oxidation prevention capability and slump resistance.
- e) Lack of nitrogen during reflow increases oxidation.

**Hypothesis 2 = Wetting does not occur in reflow**

- a) Excess BGA component warpage at temperatures in the SAC reflow range results in lack of wetting and absence of metallurgical bond.
- b) Insufficient solder paste volume or poor paste transfer efficiency.
- c) Excess solder paste hot slump. (attribute of solder paste)
- d) Inverted reflow exacerbates dewetting of BGA solder joints and increases risk of HOP.

**Hypothesis 1 – Detail**

- a) Flux in Solder Paste Burns off Prior to Wetting
  - o Soak and ramp times are the 2 key process recipe variables that impact the probability of flux burn off prior to wetting
  - o Flux dipping with paste flux + T4 SAC305 Solder Paste. 0 HOP defects on 150 trial boards. Upper bound estimate (95% confidence) for HOP defect rate is 1.98%
  - o Flux burn off is a contributing factor to HOP defects.
- b) Solder Paste-Ball Melting Point Delta
  - o Alternate parameters with SAC305 solder balls was tested with a controlled build.
  - o 100 boards were built, and parts were pried off and inspected for HOP defects.
  - o 0 defects on 100 board trial. Upper bound estimate (95% confidence) is 2.95%.
  - o Melting point delta between SAC105 solder balls and SAC305 solder paste is a contributing factor in the creation of a HoP.
  - o However, use of SAC105 balls is critical to pass OEM and JEDEC drop tests.
  - o BGA supplier conducted comparative drop-to-failure experiments and with Weibull analysis demonstrated that use of SAC305 solder balls negatively impacted drop survivability for BGA solder joints.
  - o Hence, no change was made to alloy composition for BGA solder balls.
- c) High Gold Content in Intermetallic on Surface of Component Solder Spheres
  - o An examination by XRF of solder spheres on virgin components showed a gold concentration of 0.86 +/- 0.08% for all balls examined.
  - o Industry experts said that gold concentration of this range should not result in HOP defects.
  - o Hypothesis abandoned.
- d) Solder paste formulations differ in oxidation prevention capability and slump resistance.
  - o Tests demonstrated a massive variation in HoP by switching solder paste only.
  - o Testing demonstrated that solder pastes with poor resistance to slump had higher HoP rates.
  - o See **Table 2** for a summary of paste performance on HoP.
  - o See **Table 3** in the appendix for full testing detail.
- e) Lack of nitrogen during reflow increases oxidation and poor solder risk

- Experiments were conducted using Nitrogen Gas Reflow <500ppm O<sub>2</sub>.
- Using Solder Paste Company A
- HoP Defect Rate was measured at 0% in 188 samples.  
(Upper bound estimate (95% confidence) of the true HOP defect rate was 1.58%)

**Table 2 – Ability of different solder pastes to reduce or increase HoP**

Company	Powder Type	Result
A	III	5% HOP
B		Poor Printing Performance
C	IV	Poor Printing Performance
	V	21 HOP on 100 boards
D	IV	0 HOP on 200 boards
E	IV	> 80% HOP

**Hypothesis 2 – Detail**

a) BGA Component Warpage

Hypothesized factors that impact BGA component warpage during reflow:

1. Silicon Die thickness
2. Mold Cap Compound T<sub>g</sub>
3. Mold Cap Compound CTE
4. Mold Cap Thickness

There are no issued standards with pass/fail criteria for warpage at reflow temperatures. The JEDEC standard JESD22B-112<sup>8</sup> only describes the shadow moiré method of determining warpage at elevated temperatures. The draft IEC Standard IEC 601191-6-19<sup>9</sup> “Measurement methods of package warpage at elevated temperature and the maximum permissible warpage”, describes both a shadow moiré and a laser scanning method, plus has listed pass/fail criteria.

Specific comparative experiments to test the impact of BGA warpage on HOP defect rate were not conducted. General SMT industry guidelines recommend a 100 micron max limit. This parameter is difficult to tightly control in mass production and testing for this parameter is destructive. As a result a conservative limit of +/- 50 microns max deviation from flat was specified by the OEM.

The BGA supplier has conducted structured experiments to optimize the BGA component design recipe to reduce warpage at peak reflow temperature to less than 50 microns.

b) Insufficient solder paste volume or poor paste transfer efficiency

- It was hypothesized that better transfer efficiency = better fluxing capacity
- This hypothesis was not proven as solder paste volume measurement was not used during HoP testing.

c) Excess solder paste hot slump (attribute of solder paste)

- During testing of solder paste E, it was observed that the slump performance of this solder paste appeared to be very poor.
- Solder paste E also yielded > 10x as many HoP defects during testing.

d) Inverted reflow causes dewetting of BGA solder joints and increases risk of HOP.

- XRF testing of HoP solder joints confirmed Ag concentrations to be different in both “Head” and “Pillow”. Therefore the solder of the solder paste and the solder of the solder ball never had the opportunity to co-mingle in the first place.

**Hypothesis Validation**

After careful review of the hypotheses, the following were selected for solution validation.

- Solder Paste D
- BGA Mold Cap 3C
- Modified Oven Recipe (Reduced TAL)

BGA Mold Cap 3 C warpage data was collected and Monte Carlo simulation of expected distribution for component warpage was conducted using a Crystal Ball model.

## Monte Carlo Simulation

### Monte Carlo Simulation Explained

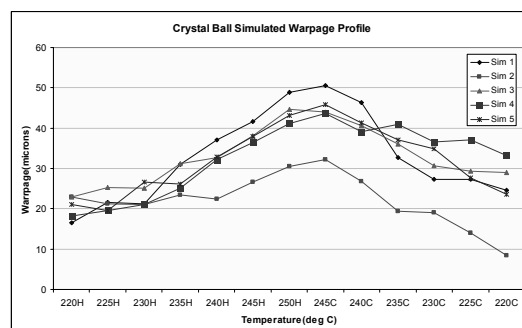
Monte Carlo simulation is a method of pre-generating a column of numbers that follow a specified distribution for a parameter or variable. For example, if the specified distribution is Normal with mean = 0 and standard deviation = 1, then using a random number seed, it is possible to generate a large column of random numbers (typically a million), that are distributed with mean = 0, standard deviation = 1, with a bell shaped histogram with percentiles conforming to the properties of the normal distribution. Warpage profile for BGA components is characterized by the warpage at 220C heating, 250C peak and 220C cooling points. If we were to actually measure the warpage profile on a million parts, we would find that no two parts generate the exact same profile. Warpage at 220C heating would follow a certain distribution with some variability. The same would hold for the warpage data for 250C peak and 220C cooling. With the data for 25 components, it would be possible to “fit” a distribution for the warpage profiles at the three critical profile points. Using Monte Carlo simulation, it would then be possible to pre-generate columns of random numbers following these pre-specified distributions, the virtual equivalent of manufacturing and characterizing a million unit high volume production run. The simulated profiles for these virtual parts could then be examined one at a time and any one that exceeds 50 microns could be tagged as a defect or at risk component. With this method, we now can estimate the fraction of at risk components that this OEM will receive when high volume manufacturing commences and the effects of multiple sources of variation begin to manifest themselves.

### Crystal Ball Explained

Crystal Ball is an add-on package to Microsoft Excel that enables us to specify distributions for variables or parameters and run Monte Carlo simulations. Crystal Ball has many more capabilities beyond Monte Carlo simulation, including sensitivity analysis and optimizers, but the decision support problem for the HOP issue did not warrant the use of any of them. Specification of distributions, limits and assumptions are done using menu buttons on toolbars within the standard MS Excel Windows GUI environment.

The warpage distributions at all profile points from 220 heating to 220C cooling indicated that they all followed a Gaussian curve and hence a mean and standard deviation could be fitted. Analysis of the profile data on the 32 parts indicated that the warpage at each profile point was correlated to the warpage at the preceding profile point. Correlation coefficients ranged from 0.7 to 0.99. Crystal Ball provides us the capability to account for this correlation between the warpage distributions at multiple profile points. **Figure 3** below captures a random selection of 5 simulated profiles from a million generated.

Monte Carlo simulation was completed for a million trials. For each trial, the maximum warpage was captured. Figure 3 shows the distribution of max warpage for the million trials. The blue area identifies the fraction of a high volume production parts that will be expected to have max warpage less then 50 microns, and the red area identifies the fraction of the population that will be at risk for HOP defects because max warpage exceeds 50 microns. The simulation predicted that 10% of the incoming BGA processors would be at some risk for HOP defects. It is important to note that there were still several unknowns regarding the relationship between max warpage and probability of a HOP defect. Hence, senior management was advised not to interpret the 10% “at risk” as equivalent to a forecast of 10% probability of HOP defects.



### Figure 3 – Crystal Ball Simulated Warpage Profile

#### Validation Testing

Validation Testing was carried out at two separate manufacturing locations over three separate builds of sample size 832. Two lab inspectors were selected from 4 tested, after attribute GR&R testing was completed. All 2496 samples were inspected using pry off and microscopic inspection of BGA solder joints (**Table 3**).

#### Test Vehicle Attributes

- PCB = Final – Cu Balanced
- BGA = DOE 3C, Mechanical samples FBGA.
- RF Shield = Final

#### Reflow Recipe Factor Description

- Nitrogen = 500pm O2
- Additional addition of dipped Flux = A (Y = yes, N = no)
- Paste = D

**Table 3 – Validation Testing Results**

Validation Run	Mfg Line	Sample Size	Gas	Flux	Paste	Samples HoP
1	1	832	Air	Y	D	0
2	3	832	Air	N	D	0
3	1	832	Air	N	D	0

#### Other Validation Run Details

- PCB lot code was randomized between runs.
- Current Qualified Production Reflow Profile was used.
- 4 mil E form stencil, no stepping.
- Print speed was set at 2” per second.
- Print pressure was controlled per established control plan.
- Ambient humidity and temperature was not monitored.
- No unusual events or stoppage of flow was observed during the validation runs.
- Solder paste lot code was not changed during a build.

#### Process Robustness Testing

To evaluate how robust the validated solution was, further experimentation was carried out and some variables were modified to deliberately increase the risk of HoP. See **Table 4**.

**Table 4 – Process Robustness Testing Results**

Validation Run	Mfg Line	Sample Size	Gas	Flux	Paste	Samples HoP
1	1	100	Air	N	D	0

#### Test Vehicle Attributes

- PCB = Final – Non Cu Balanced

- BGA = 1mm Dev, Mechanical samples FBGA. (Lot Measured Warpage > 90 microns)
- RF Shield = Final

#### Reflow Recipe Factor Description

Nitrogen = 500pm O2

- Additional addition of dipped Flux = A (Y = yes, N = no)
- Paste = D

#### Other Validation Run Details

- PCB lot code was randomized between runs.
- Current Qualified Production Reflow Profile was used.
- 3 mil E form stencil, no stepping. (Simulate Lower Paste Volume)
- Print speed was set at 2” per second.
- Print pressure was controlled per established control plan.
- Ambient humidity and temperature was not monitored.
- No unusual events or stoppage of flow was observed during the validation runs.
- Solder paste lot code was not changed during a build.
- 

#### **Conclusions**

HoP is a challenging and complex soldering problem with much risk for the OEM. For the OEM, the lack of a reliable test method in mass production leaves the OEM with no choice but to find and test a solution to the point that it can be assumed the HoP is no longer a problem.

From the data collected, it is clear that any mechanism which alters surface oxidation of the BGA ball results in significant variation in the HoP incidence.

The addition of flux dipping, and/or N2 gas reflow both reduced the HoP defect rate to zero. (For practical purposes, zero equals an upper bound estimate (95% confidence) of 0.1% for the true HOP defect rate.

The switching of solder paste also demonstrated a huge affect on the HoP incidence. The testing of Pb-free solder paste from five different 1<sup>st</sup> Tier suppliers showed that whilst one paste virtually eliminated HoP, another made it over 10 times worse than the original defect rate.

Finally, it would seem that a solder paste with poor slump performance does not do so well in the fight against HoP. This OEM’s testing indicated that the worse the slump, the greater the incidence of HoP.

#### **Acknowledgements**

The authors would like to thank our executive sponsors, Benson Tendler, Vice President of Hardware Quality and Fintan Doyle, Senior Director of New Product Introduction for support of this work. We would also like to thank Linda Galvis for the moiré work and Madusha Cooray and Ray Kodali for carrying out all the BGA pry and examination work.

#### **References**

1. “Pb-free: Fact or Fiction?”, <http://www.circuitsassembly.com/cms/news/6458>, April 18, 2008.
2. Karl Seelig, “HIP Defects in BGAs”, Circuits Assembly, pp 28-31, December 2008.
3. Tim Jenson, “The Graping Phenomenon: Improving Pb-Free Solder Coalescence through Process Optimization and Materials” Proceedings of APEX 2008, Las Vegas.
4. Chrys Shea, “Step the HOP”, p 33, Circuits Assembly, August 2008.
5. Chrys Shea, “HOP-ing Mad”, Circuits Assembly, pp 72-73, July 2008.
6. “Koki No-clean Lead Free Solder Paste Anti-Pillow Defect S3X58-M406-3 series Product information”, version 42016e, August 29, 2006, [www.ko-ki.co.jp](http://www.ko-ki.co.jp)
7. Rick Lathrop, “BGA Coplanarity Reduction During the Ball Attach Process”, Capital SMTA meeting, June 5, 2007.
8. JESD22B-112, “High Temperature Package Warpage Measurement Methodology”, August 2005.
9. IEC 601191-6-19 (draft), “Measurement methods of package warpage at elevated temperature and the maximum permissible warpage”



**Appendix 1 Table of Experimental Results**

Exp' #	Mfg Line	BGA		Printer	Solder Paste			Reflow Recipe				Expt. Sample Size	# of HoP	p Value
		Mold Cap	Alloy		Paste	Alloy	Flux Dip	Oven	Gas	Pk Temp	TAL			
0	1	Imm Dev	SAC105	A	A	SAC305	No	1	Air	242	70	100	5	-
1	1	Imm Dev	SAC105	A	C	SAC305	No	1	Air	242	70	100	21	0.005
2	1	Imm Dev	SAC105	A	E	SAC305	No	1	Air	242	70	100	85	0
3	1	Imm Dev	SAC105	A	D	SAC305	No	1	Air	242	70	200	10	0.5
4	1	Imm Dev	SAC305	A	D	SAC305	No	1	Air	242	70	52	3	0.42
5	2	Imm Dev	SAC105	B	A	SAC305	No	2	N2	242	70	100	0	0.012
6	2	Imm Dev	SAC105	B	A	SAC305	No	2	N2	242	70	88	0	0.017
7	1	Imm Dev	SAC105	A	A	SAC305	Yes	1	Air	242	70	56	0	0.046
8	2	Imm Dev	SAC105	B	D	SAC305	No	2	N2	242	70	96	0	0.014
9	2	Imm Dev	SAC105	B	D	SAC305	No	2	Air	242	70	40	0	0.076
10	3	Imm Dev	SAC105	A	D	SAC305	Yes	3	Air	242	70	150	0	0.003
11	1	Imm Dev	SAC105	A	D	SAC305	Yes	1	Air	242	70	152	0	0.003
12	1	Imm Dev	SAC105	A	D	SAC305	No	1	Air	242	70	148	0	0.003
13	1	Imm Dev	SAC305	A	D	SAC305	No	1	Air	242	70	100	0	0.012
14	1	DOE 3C	SAC105	A	D	SAC305	No	1	Air	242	70	148	0	0.003

**P Values -**  
 - Statistically significant increase from base line OR no significant decrease from baseline.  
 - Insufficient data to determine if statistically significant decrease from baseline.  
 - Statistically significant decrease from baseline.  
 □ □ = 5%, Desired Power = 80% min to detect a 3% drop from baseline.