J. Am. Ceram. Soc., 80 [4] 831–38 (1997)

Interlaboratory Verification of Silicon Nitride Tensile Creep Properties

William E. Luecke* and Sheldon M. Wiederhorn*

Materials Science and Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Five laboratories tested NIST-supplied, pin-loaded, 76-mm-long tensile creep specimens at 1400°C under a 150 MPa load using flag-based, laser extensometry. The laboratories reported failure time and strain and supplied the individual creep curves. Only one of the laboratories produced failure times that were significantly less than the others. It is likely that their reduced failure times resulted from small load calibration and test temperature errors. After steps were taken to ameliorate these problems, three additional tests yielded failure times that agreed with those of the other four laboratories. Although the times to failure from the four laboratories that initially agreed were statistically indistinguishable, their creep curves exhibited subtle differences. These differences probably arose because the laboratories used different gage length definitions. When we recalculated the creep curves to the same gage length definition, the differences between the four laboratories whose times to failure agreed, vanished. Although a number of the specimens exhibited edge chips, creep cracks, and obvious chemical interactions with the flags, the presence of these defects did not reduce the time or strain to failure. Two additional creep tests in our laboratory, using specimens that were grossly misaligned, yielded failure times and strains that were commensurate with those from well-aligned specimens.

I. Introduction

In the last decade, independent laboratories have produced several databases of high-temperature creep rupture properties of commercial silicon nitrides. Often these databases report very different times to failure (or minimum creep rates or failure strains) for nominally identical materials.[1] Do the differences arise from variations in the materials that each laboratory tested, or do they arise from differences in testing procedure between the laboratories? There has been evidence that differences in material properties could be the source of some of the variability, since replicate tests within a laboratory often produce five-fold scatter in properties (for an example, see Ref. 2). There has never been, however, a systematic study to determine how the reported tensile creep properties of a silicon nitride differ between laboratories.

Recently we obtained a quantity of Si_3N_4[3] (SN88, NGK Insulators, Nagoya, Japan)† that exhibits highly repeatable creep properties. Unlike other silicon nitrides, the standard deviation of the times to failure of this material for five replicated tests, conducted in our laboratory, is about 10% of the mean value.[4] The excellent repeatability of this material encouraged us to organize an interlaboratory study of creep testing to quantify the differences between laboratories, because it minimizes the number of specimens necessary to observe the expected small differences between laboratories. The larger variability of other silicon nitrides would have required unacceptably large numbers of specimens.

The primary goal of this informal interlaboratory study was to enable the participants to identify ways in which their testing procedures may give measurements that differ from those that other laboratories obtain. A secondary goal was to quantify the differences that might be expected in creep properties measured in different laboratories. We hope that the results will be of interest to other laboratories, however. Generally, this study followed the procedure outlined in ASTM E 691[5] (Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method). It deviates from it in that there are only five participating laboratories, rather than the minimum of six recommended.

Five laboratories out of seven invited (Table I) agreed to participate. Although other laboratories do high-temperature creep testing, a necessary and unfortunate limitation of this study was that the participating laboratories be able to test the already-fabricated NIST specimens. This requirement excluded laboratories that normally test buttonhead or shoulder-loaded specimens. One of the declining laboratories did so because they discovered that they could not test the specimen supplied. The other, though physically able to test the specimen, had obligations that made it impossible to complete the testing in the allotted time.

II. Testing Procedure

A necessary requirement of the tests was that the laboratories be able to complete the entire test cycle in less than several weeks. Although more replication would allow better statistics and perhaps might reveal other interesting differences, requiring more tests would have made it impossible for most laboratories to participate. With this constraint in mind, the instructions asked the laboratories to test at least three of five specimens to failure under a single condition: 1400°C, 150 MPa. Extensive testing at these conditions in our laboratory[4] ensured that each test would take less than 100 h.

The instructions to the participants strove to prescribe as little of the testing procedure as possible. Before applying the full test load, the laboratories were to anneal the specimens in the creep rig under a 10 MPa load for 24 h, even though this procedure might not be part of their testing protocol. After loading, they were to monitor the strain by their own technique until failure. In particular, there were no restrictions on the gage length they were to use. After completing the testing, they were to return the broken specimens, a brief report on their testing protocol, the data files containing the individual creep (strain vs time) curves, and completed data sheets summarizing their observations. On the data sheet for each specimen, they were to report the failure time and strain, the gage length employed, the location along the gage length of the final fracture, and an interpretation of the fracture surface.

All five laboratories use very similar flag-based, laser-extensometry systems for creep measurement. A recent publication from our laboratory details and analyzes the errors and

John J. Petrovic—contributing editor

Manuscript No. 191688. Received July 9, 1996; approved September 24, 1996.

Supported by the Ceramic Technology Project, DOE Office of Transportation Technologies, under Contract No. DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

*Member, American Ceramic Society.

†The use of commercial designations or company names is for identification only and does not indicate endorsement by the National Institute of Standards and Technology.

Table I. Participants in This Interlaboratory Study

Laboratory†	Experimenter
National Institute of Standards and Technology, Materials Science and Engineering Laboratory	William Luecke
AlliedSignal Technology	Chien-Wei Li/Charles Gasdaska
Oak Ridge National Laboratory, Metals and Ceramics Division	H. T. Lin
Lehigh University, Department of Materials Science	Junghyun Cho
St. Gobain Industrial Ceramics, Northboro Research and Development Center	Jon Wade

†The order of the laboratories listed may not correspond to their numerical designation.

uncertainties inherent to flag-based extensometry.[6] Figure 1 is a schematic diagram of the specimen,[4] designated as SR76, used in the testing. We developed this specimen by modifying an original design from Wade *et al*.[7] It has a segment of uniform cross section that is 19 mm long, and the 19 mm radius of curvature from the head to the gage length results in only a 3.7% elastic stress concentration in the transition region. We have successfully conducted room-temperature strength tests up to 700 MPa using this specimen.

Because the SR76 was not the usual specimen for any of the other laboratories, we found it necessary to supply some of the laboratories with loading pins and flags. We sent flags to all the laboratories, except laboratory 7, as well as loading pins to laboratory 1. The flags we supplied are different in geometry (see Fig. 2) than those the laboratories usually employ. In most cases, their flags would not have clipped on to the NIST SR76 specimen.

III. Results and Discussion

Table II displays the results supplied by the different laboratories, arranged by laboratory. For each laboratory, the table shows the failure time, t_f; failure strain, ε_f; the location of the failure, measured from the geometric center of the specimen; the gage length used to calculate the strain; and a column indicating whether the specimen failed inside or outside the uniform cross section of the specimen. Figures 3(a–e) show the creep curves for all the specimens tested.

Both Table II and Fig. 3 show that times to failure from laboratory 6 are much shorter than those from the other laboratories. In addition, whereas 13 of 15 creep curves from laboratories 1, 3, 4, and 7 show tertiary creep, only two of five creep curves from laboratory 6 do. Figure 3 also shows that the difference between creep curves from a given laboratory is less than the difference between curves from different laboratories. In a later section, we shall attempt to quantify these differences. No conclusions should be drawn from the noise level of the readings in each curve, as we averaged them to facilitate manipulation. We adjusted one creep curve from laboratory 4, after consultation with the experimenter, by splicing two sections together, but aside from this, Fig. 3 shows the strain-vs -time curves as the laboratories supplied them.

From the creep curves, the participants extracted two parameters: failure time, t_f, and failure strain, ε_f. Although laboratories often report minimum strain rate, it is difficult to interpret differences in this parameter, because it is sensitive to both the time to failure and the strain measurement technique. For this reason we will not consider differences in the minimum strain rate. To facilitate comparison of the underlying creep behavior, we took the creep curves and extracted the strain at a fixed time, which we chose as the time of failure of the shortest-lived specimen, $\varepsilon_{t_{min}}$, of the entire group of 20 creep curves.

(1) Time to Failure

Because the failure time, t_f, can be extracted directly from the test data, it is the most objective indicator of differences in test technique between laboratories. Typically, the apparent activation energy for failure time of silicon nitride is −1000 kJ/mol, and the stress dependence ($t_f \propto \sigma^{-m}$) is often greater than 10. Because of this extreme sensitivity, small deviations in test temperature and load can produce large deviations in failure time. For some failure modes, large deviations can indicate poor alignment of the load train as well.[8] While these sources may also influence the strain to failure and minimum strain rate, interpretation of these latter two is complicated by uncertainties in gage length measurement and definition. Figure 4 shows that the failure times of four of the five laboratories agree. An analysis of variance of the failure times confirms what Fig. 4 shows: the failure times for laboratory 6 are significantly less than those of the other laboratories.

What might be the origin of the discrepancy between laboratory 6 and the rest of the laboratories? The fracture surfaces of their specimens did not appear any different from those of the other laboratories. In addition, the failure locations along the gage length were distributed similarly to those of the other laboratories. Finally, the fact that the strains to failure are similar to those obtained in the other laboratories makes premature failure, such as might be induced by misalignment, less plausible as an explanation. These observations and the known sensitivity of the failure time on temperature and stress both suggest that the actual test conditions in laboratory 6 deviated from the desired conditions.

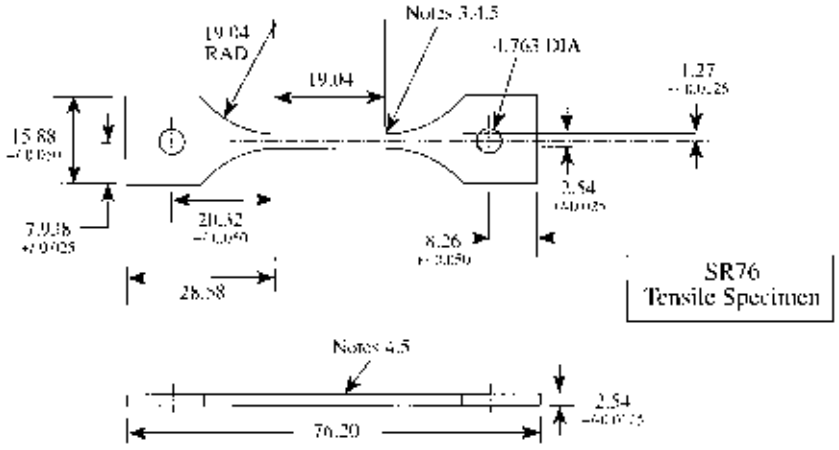

Fig. 1. Specimen used in the creep testing.

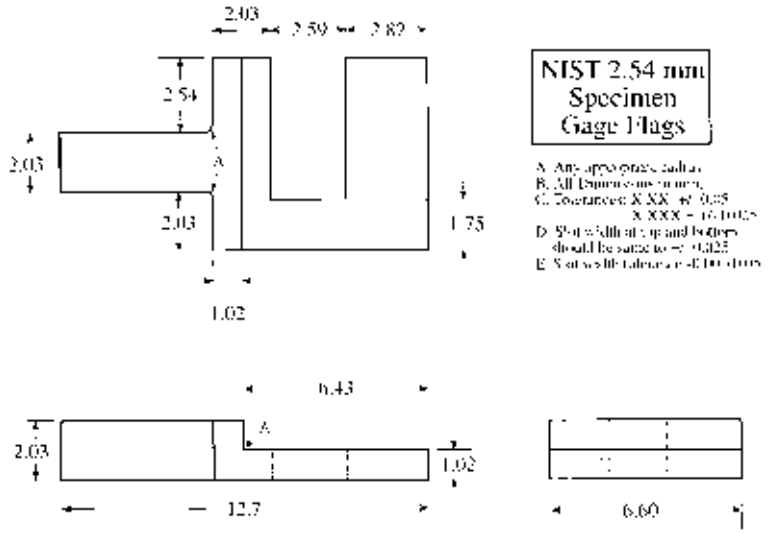

Fig. 2. Flag supplied to laboratories 1, 3, 4, and 6.

Table II. Results for All Specimens Tested

Lab	t_f (h)	ε_f	Failure location* (mm)	Gage length (mm)	Failure location with respect to uniform cross section
1	81.5	0.0324	0.70	14.34	Inside
1	76.7	0.0307	6.20	14.88	Inside
1	74.7	0.0285	−2.00	17.03	Inside
1	58.5	0.0229	0.90	16.45	Inside
3	86.7	0.0283	8.97	11.81	Inside
3	83.6	0.0255	4.10	13.57	Inside
3	82.0		3.87		Inside†
3	72.5	0.0294	−7.70	12.31	Inside
4	88.2	0.0288	−1.40	19.20	Inside‡
4	74.6	0.0230	−8.40	18.10	Inside
4	73.0	0.0260	1.60	17.20	Inside§
4	72.5	0.0240	0.60	17.50	Inside¶
6	46.7	0.0276	0.21	10.83	Inside
6	40.8	0.0265	−9.96	14.77	Outside
6	37.1	0.0273	−5.00	10.32	Inside
6	36.7	0.0252	−8.79	11.86	Inside
6	36.4	0.0199	−7.70	9.96	Inside
7	74.8	0.0332	−0.42	15.60	Inside
7	74.0	0.0358	10.00	15.67	Outside
7	73.4	0.0325	3.25	14.35	Inside
7	71.8	0.0273	−10.65	15.08	Outside
			Mean values		
1	72.8	0.0286	1.45	15.67	
3	81.2	0.0277	2.31	12.56	
4	77.1	0.0255	−1.90	18.00	
6	39.6	0.0253	−6.25	11.55	
7	73.5	0.0322	0.55	15.18	
			Grand mean		
	67.4	0.0277	39.13		
			Special tests		
			Misaligned specimens		
3	99.0	0.0287	−8.74	12.36	Inside
3	67.5	0.0194	−6.74	11.72	Inside
			Laboratory 6 retests		
6	101.2	0.0294		12.92	
6	82.4	0.0356		8.65	
6	61.4	0.0265		9.55	

*Measured from the midpoint of the specimen. The uniform cross section lies between −9.52 mm and 9.52 mm. †Data collection failed—no strain measurements. ‡Experiment contains a 3 h segment at reduced load due to load control failure. §Specimen was reloaded after 0.5 h at load due to loading pin failure. ¶Specimen was reloaded after 3 h at load due to loading pin failure.

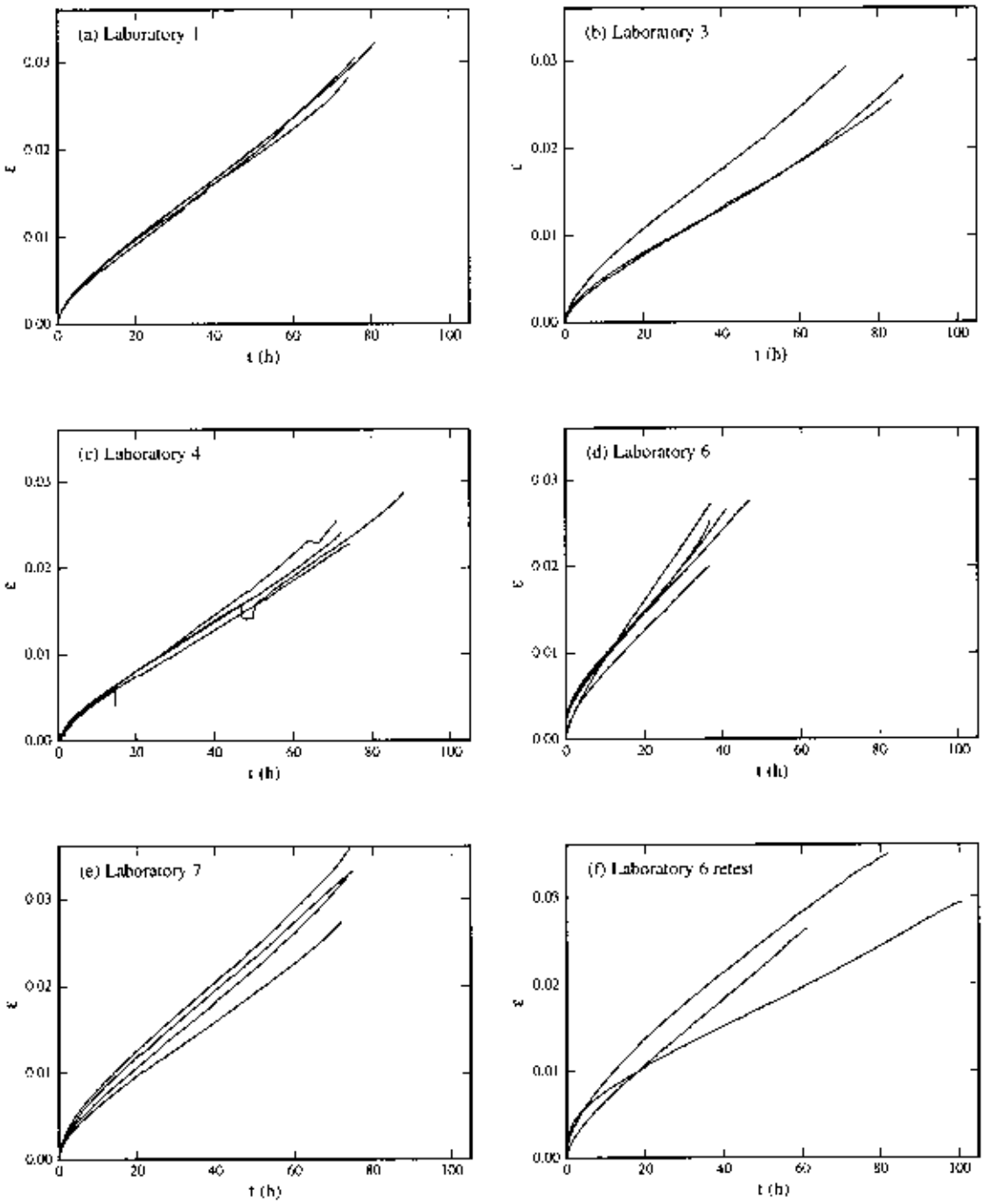

Fig. 3. Creep curves for all the specimens tested: (a) laboratory 1; (b) laboratory 3; (c) laboratory 4; (d) laboratory 6; (e) laboratory 7; (f) laboratory 6, second set of tests after discussing origin of time to failure discrepancy.

Using unpublished work in our laboratory, we have developed an expression for the time to failure of this silicon nitride as a function of temperature and stress over a wide range of test conditions. In the neighborhood of the test conditions (1400°C, 150 MPa) the failure time decreases 3.22 h per 1°C increase in temperature. The corresponding derivative of failure time with respect to stress is similarly large: −2.92 h/MPa. After informing laboratory 6 of their discrepancy in failure time, we collaborated to locate its origin. To verify their test temperature, they used the same NIST-calibrated, Type B (Pt–30% Rh vs Pt–6% Rh) thermocouple we used to verify our temperatures. Using this thermocouple, as well as their own unshielded, Type S (Pt–10% Rh vs Pt) thermocouples and a dummy specimen, they discovered that the test temperature was 1406.9°C instead of the desired 1400°C. In analyzing their load train, they also discovered a 24.4 N (3.8 MPa) error in the applied load. Using the expression for failure time, the expected difference under these conditions between laboratory 6 and the grand mean is

$$t_f(1406.9°C, 153.8\ \text{MPa}) - t_f(1400°C, 150\ \text{MPa}) = 28.6\ \text{h} \quad (1)$$

which compares well with the observed 36.5 h difference.

After identifying these small deviations as possible sources for the failure time discrepancy, we supplied laboratory 6 with an additional three specimens to verify that the changes in test protocol they implemented were successful. For this round of testing, however, laboratory 6 used flags of its own design. In contrast to the 2.03 mm × 2.03 mm dimensions of the arm of the NIST flag (see Fig. 2), these flags were 2.54 mm high and 5.08 mm deep. Furthermore, the region of the flag that clips to the specimen is the full 2.54 mm height, rather than the 1 mm

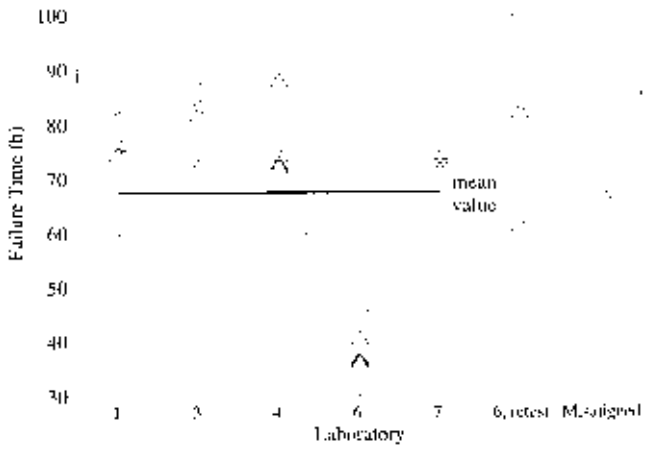

Fig. 4. Failure times measured by the five laboratories. The solid line is the mean value for all the laboratories, while the dashed lines are the 95% confidence interval for the mean. Also included are the failure times for the tests laboratory 6 conducted after we discussed possible origins of their different failure times. The figure also shows data for the two misaligned specimens that we tested.

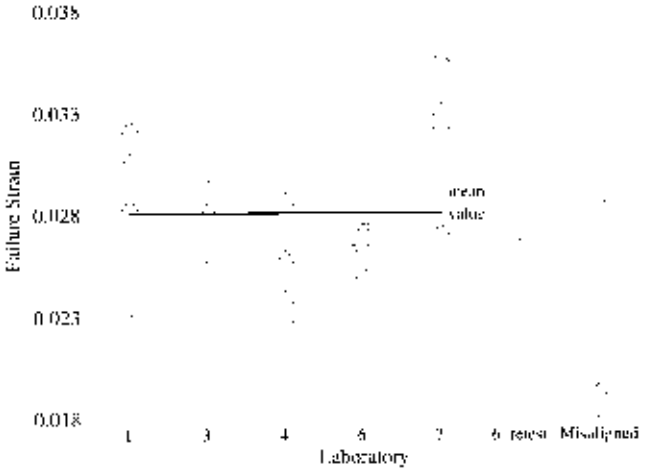

Fig. 5. Strains to failure measured by the five laboratories. The solid line is the mean value for all the laboratories, while the dashed lines are the 95% confidence interval for the mean. Also shown are the strains to failure for the second set of tests that laboratory 6 conducted, as well as those for the two misaligned specimens.

dimension of the NIST flag. Figure 3(f) shows the three resulting creep curves, and Table II and Fig. 4 summarize the relevant information. An analysis of variance of the failure times, after substituting this new data for the original data of laboratory 6, indicates no difference between the laboratories at the 95% significance level. The changes were successful, at least in terms of their effect on failure time. However, the scatter in the creep curves of the second set is larger, and only one of the three curves exhibits tertiary creep.

(2) Strain to Failure

Strain to failure is a second useful diagnostic parameter, but identifying differences between laboratories is complicated both by differences in test technique and by differences in interpretation of the extension data. Figure 5 shows the strains to failure arranged by laboratory. An analysis of variance indicates no difference in strain to failure between the laboratories at the 95% significance level. Replacing the original data from laboratory 6 with the three data from the second set of tests and conducting the analysis of variance again does not change this conclusion.

(3) Strain at Minimum Failure Time

Figure 3 shows that the creep curves obtained by the five laboratories fall into distinct bands, even though the analysis of variance of the failure strains indicates no statistically significant difference. The reason for the latter observation is that the strain to failure is controlled by both the underlying creep mechanism and the failure mechanism. Consider the creep curves of laboratory 1 (Fig. 3(a)). Certainly, the creep behaviors of the four specimens are nearly identical, but the strains to failure differ because one of the specimens failed much earlier than the other three. To examine the difference between the measured creep curves, it is useful to examine the strain at a fixed time before failure, rather than at the failure time. Figure 6 shows the strains at 36.4 h, the time of failure of the shortest-lived specimen (from the initial tests by laboratory 6), which is about one-half the mean lifetime of the rest of the specimens. The mean value includes the values from the second set of tests of laboratory 6. It is possible to use the results of an analysis of variance to look for differences between the strains, using Tukey's approximate multiple comparisons procedure for unequal sample sizes.[9] That procedure indicates that there is evidence that the creep curves fall into two groups: laboratories 7, 6 (second set), 1, and 3 form one group, and laboratories 1, 3, and 4 form the other. Laboratories 1 and 3 are common to both groups.

One possible explanation for this grouping is that laboratories 1 and 7, which are in the group with the larger characteristic strains, both define the gage length as the distance from the top of the bottom flag to the bottom of the top flag. Laboratories 3 and 4 actually compute two strains for each experiment: one where the gage length is defined as the distance from the bottom of the bottom flag to the bottom of the top flag and the other where it is defined as the distance from the top of the bottom flag to the top of the top flag. The reported creep curve is the average of these two curves. Laboratory 6 defines the gage length as the distance from the bottom of the bottom flag to the top of the top flag. Figure 7 illustrates these definitions.

Setting aside the question of the rectitude of these different gage length definitions, the fact that they exist means that the different laboratories will report different strains in analyzing *the same* test. As all the laboratories can measure the change in gage length quite accurately, we may assume that the changes in gage length are correct. The nominally 15-mm gage lengths that laboratories 1 and 7 employed correspond to 17-mm gage lengths if measured in the systems from laboratories 3 and 4,

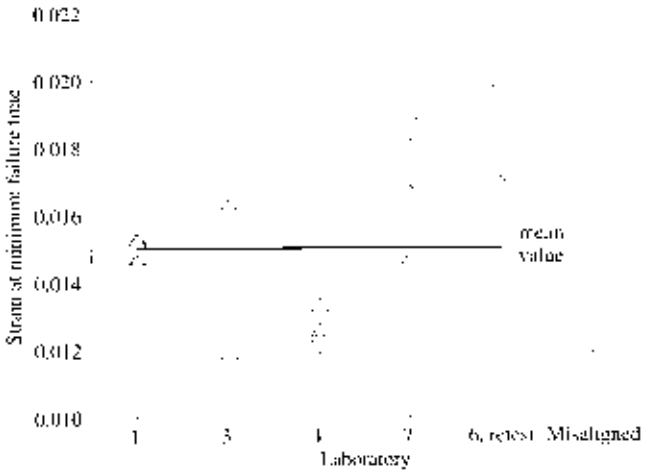

Fig. 6. Strain at the failure time of the shortest-lived specimen (t_f = 36.4 h). The solid line is the mean value for laboratories 1, 3, 4, and 7, while the dashed lines are the 95% confidence interval for the mean. Data from laboratory 6 are excluded from the calculation. The figure also shows the strains for the second set of tests laboratory 6 conducted, as well as for the two misaligned specimens.

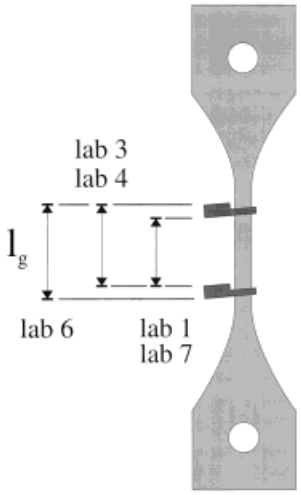

Fig. 7. The three gage length definitions used by the participating laboratories. Note that the orientation of the flags with respect to the specimen may not be the one the laboratories actually employed.

since all the laboratories used flags that were 2 mm high. Therefore, all other variables being equal, the strains that laboratories 1 and 7 measured should exceed those of laboratories 3 and 4 by about (15 mm + 2 mm)/15 mm = 1.13. By the same token, laboratory 6 reports a nominal 11-mm gage length. To correspond to laboratories 3 and 4, it is necessary to divide their strains by (11 mm − 2.54 mm)/11 mm = 0.77, because they used their own 2.54 mm thick flags for the second set of tests. Rescaling the creep curves, using the actual gage lengths from Table II, and conducting a second analysis of variance shows that only laboratory 6 differs from the other laboratories. Figure 8 shows the creep curves after rescaling.

Although rescaling the creep curves to a common gage length definition harmonized four of the five sets of creep curves, it drove those from laboratory 6 further from the group. In addition, the scatter of the second set of tests from laboratory 6 is much larger than from the first set. We do not understand the reason for this deviation.

(4) Failure Origins

All of the specimens failed by the growth of a crack that originated at one of the four corners of the gage section. The failure sites were evenly distributed along the length of uniform cross section of the 21 specimens. The surfaces of the cracks

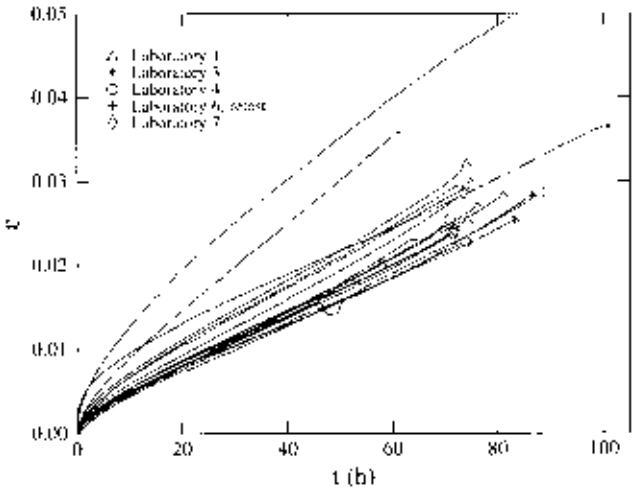

Fig. 8. Creep curves for laboratories 1, 3, 4, 7, and 6 (second set of tests), with those for laboratories 1, 6, and 7 rescaled to correspond with the strain measurement of method of laboratories 3 and 4.

were always quite rough, in contrast to the much smoother final failure. Apparently, these corner cracks grew for some time before rapidly propagating to failure. Morphologically, they are similar to the cracks that Menon *et al*.,[2] Luecke *et al*.,[10] and Wereszczak et al.[11–13] have observed in other silicon nitrides. In their earlier works,[11,12] Wereszczak *et al*. refer to these as stress-corrosion cracking zones, but in a later work[13] they show that they appear in specimens tested in argon as well as air. Most of the specimens tested contained an additional one to five of these creep cracks along the uniform cross section. The number of creep cracks appearing per specimen correlated weakly with increasing failure time, excluding data from laboratory 6. There was no evidence that any laboratory produced more creep cracks per specimen than any other, and the number did not correlate with strain to failure. The creep cracks do not seem to be the origin of the tertiary creep, since there was no correlation between the number of creep cracks in a specimen and the amount of tertiary creep. Furthermore, all the specimens that laboratory 6 tested show creep cracks, but only two of their creep tests show tertiary creep.

There are two possibilities for the origin of the creep cracks. They may have been present in the specimen at the start of the experiment and grew slowly during the test. A second possibility is that the creep process itself, which usually involves extensive cavitation of the silicate interstitial pockets,[2,10,14] nucleates the life-limiting crack. Environmental attack may accelerate the growth rate of the crack. The corners of the gage sections of 9 of the 21 specimens tested were chipped during manufacture. Although these chips are in the same area from which the life-limiting cracks emanate, the mean failure time and strain (excluding laboratory 6) for chipped specimens were statistically indistinguishable from those for the unchipped specimens. Furthermore, unpublished tests in our laboratory, using 51-mm specimens of the same material, showed that transverse grinding of the gage length with a 240 grit, resin-bonded wheel prior to testing does not reduce the failure time of the material. Preexisting damage in the specimen, therefore, is not a likely candidate for the origin of the life-limiting flaws.

(5) Effects of Specimen Misalignment

Researchers generally agree that the alignment of the specimen in the testing frame is central to the production of accurate results in tensile testing. For instance, the recently adopted ASTM standard for creep testing of ceramics (C 1291-95)[15] calls for not more than 5% bending during the test. In the specimen of this study, when the loading holes are displaced symmetrically by more than 0.01 mm (10 μm) from the gage centerline, the bending strain calculated by elastic beam theory exceeds this requirement. To meet this requirement, the specification for the test specimen directed that the loading line, defined by the line connecting the centers of the loading holes, and the centerline of the gage length be aligned to within 0.021 mm (21 μm). Of course, misalignment of the loading train and the couplings may also introduce bending stresses, but verification using a strain-gaged specimen indicated that it is possible to meet the requirements of ASTM C 1291-95 in our loading train.

Generally, the specimens of this study met that rather strict tolerance. Two specimens that did not allowed us to assess the effects of specimen misalignment on lifetime and strain measurement. Figure 9 shows the traces of the centerlines of these two specimens, measured (Model BHN710 Coordinate Measuring Machine, Mitutoyu, Tokyo, Japan) before and after creep testing. The line connecting the centers of the two holes and its perpendicular bisector define the origin of the coordinate system. In one specimen, the loading holes are displaced more than 0.075 mm (75 μm) from the gage centerline. Both specimens would certainly violate the stringent alignment requirements of ASTM C-1291-95, at least with respect to elastic bending.

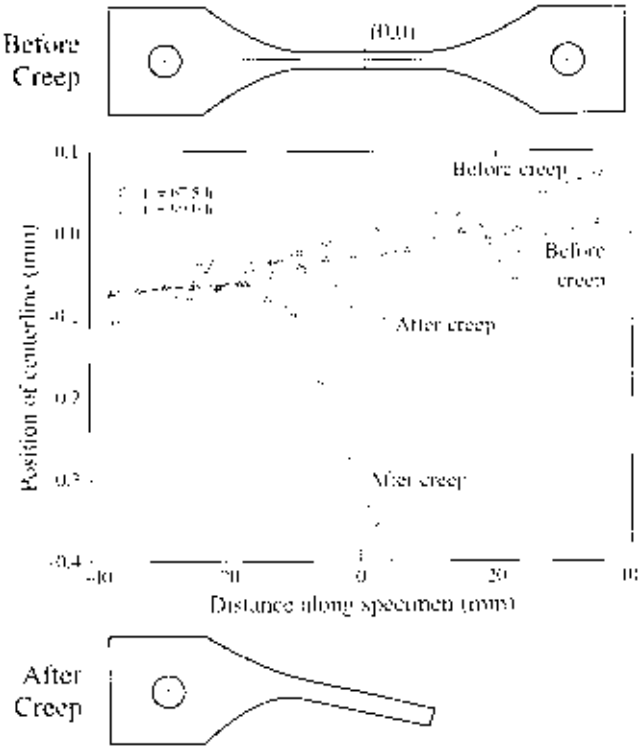

Fig. 9. Traces of the centerlines of the two misaligned specimens before and after creep. The schematic of the specimen above is drawn on the same scale as the *x*-axis of the figure and indicates the *x*-position of the loading holes. The line joining the centers of the loading holes and its perpendicular bisector define the origin of the coordinate system.

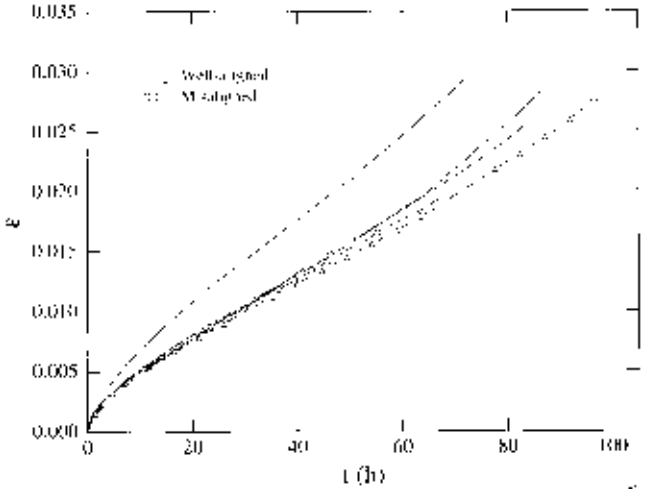

Fig. 10. Creep curves from three well-aligned specimens and the two poorly aligned specimens. Symbols on the creep curves from the poorly aligned specimens correspond to those in Fig. 9.

We tested both of these specimens under the conditions of this study. To facilitate comparison, we identify ourselves as laboratory 3. Figure 10 shows the three creep curves for the well-aligned specimens (from Fig. 3(b)) along with the two creep curves from the poorly aligned specimens. Data for the two tests also appear in Table II and Figs. 4–6. The failure times of the poorly aligned specimens bracket the range of times from the well-aligned specimens. The shorter-lived specimen does not show tertiary creep. The underlying creep behavior of the two poorly aligned specimens is clearly indistinguishable from that of the well-aligned specimens. Figure 9 also shows the traces of the two specimen centerlines after failure. As we have observed with other silicon nitrides,[6] creep has accommodated the misalignment all at one end of the uniform cross section, near the specimen neck, rather than distributing it uniformly along the gage length. This type of deformation introduces negligible errors in strain measurement.[6] Apparently, it does not significantly reduce rupture lifetime, either. The results of this small study show that specimens may be rather grossly misaligned without inducing serious error in strain or lifetime measurement.

IV. Implications for Life Prediction

One goal of testing commercial materials is to build data bases of creep–rupture properties to allow designers to predict rupture life of actual components. Uncertainties in the values of rupture life from such a database result in uncertainties about the reliability of components and can arise from several different sources. The lifetime of the component may differ from the predicted value because of subtle differences in chemistry or microstructure between the material used in the data base generation and that used for component fabrication. This might be termed the "vintage effect." Differences in lifetime might also arise from the difference in geometry between the test specimen used to gather the data and the intended application. Because the final failure of the test specimen results from the growth of a crack under uniaxial stress, and the component presumably undergoes multiaxial loading, the failure modes may be different. Until better microstructural models of creep deformation of silicon nitride appear, this problem will remain. Thirdly, there may be differences that originate from the care with which the database-producing laboratory collected the data. This study attempts to address these uncertainties.

For the silicon nitride of this study, there is very little published information to assess the vintage effect. In 1993, Watanabe *et al.*[3] reported a lifetime of 140 h for a single specimen tested at 1400°C and 150 MPa. Unfortunately, they conducted their tests on shoulder-loaded specimens, so the increased lifetime relative to that in this study could arise from a mixture of both interlaboratory differences and specimen geometry effects, as well as vintage differences. Unpublished research indicates that the vintage effect may be stronger than this comparison indicates, in that other laboratories have measured failure times for SN88 that differ by up to factors of three. Here again, though, these data come from a variety of specimen geometries and test techniques, as well as from different vintages of material. If we regard the differences as arising wholly from vintage, though, the uncertainty due to vintage probably outweighs that due to interlaboratory differences. Designers can then have reasonable confidence that data generated from reputable laboratories on a specific vintage material are correct. Clearly, a larger interlaboratory study, incorporating different specimen geometries and test techniques, is necessary to validate this supposition.

V. Conclusions

(1) Different laboratories are capable of good agreement of measured creep properties when they test material with highly repeatable properties.

(2) The difference in failure time initially obtained by laboratory 6 resulted from the combination of small temperature and stress excesses. After laboratory 6 corrected these small deviations, a second round of testing produced results that agreed with the other laboratories.

(3) The small differences between the creep curves obtained by the laboratories resulted primarily from their different definitions of the gage length.

(4) Tests of two misaligned specimens produced times and strains to failure, as well as creep curve shapes, that were indistinguishable from those produced by well-aligned specimens. This agreement indicates that excellent specimen alignment may not be necessary for routine creep testing.

Acknowledgments: We thank all the participants for their time and patience. We thank Robert Wantz for measuring the positions of the centerlines of the misaligned specimens. J.D. French drew Figs. 1 and 2.

References

[1]R. F. Krause Jr., W. E. Luecke, and S. M. Wiederhorn, "Comparison of Tensile Creep Measurements on a Hot Isostatically Pressed Silicon Nitride," submitted to *J. Am. Ceram. Soc*., 1996.

[2]M. N. Menon, H. T. Fang, D. C. Wu, M. G. Jenkins, M. K. Ferber, K. L. More, C. R. Hubbard, and T. A. Nolan, "Creep and Stress Rupture Behavior of an Advanced Silicon Nitride: Part I, Experimental Observations," *J. Am. Ceram. Soc*., **77**, 1217–27 (1994).

[3]K. Watanabe, M. Masuda, T. Ozawa, M. Matsui, and K. Matsuhiro, "Research and Development of Ceramic Turbine Wheels," *J. Eng. Gas Turbines Power*, **115**, 36–41 (1993).

[4]J. D. French and S. M. Wiederhorn, "Tensile Specimens from Ceramic Components," *J. Am. Ceram. Soc*., **79** [2] 550–52 (1996).

[5]"Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method," E 691-92, 1992 Annual Book of ASTM Standards, Vol 14.02. American Society for Testing and Materials, Philadelphia, PA, 1992.

[6]W. E. Luecke and J. D. French, "Sources of Strain Measurement Error in Flag-Based Extensometry," *J. Am. Ceram. Soc*., **79** [6] 1617–26 (1996).

[7]J. A. Wade, C. S. White, and F. J. Wu, "Predicting Creep Behavior of Silicon Nitride Components Using Finite Element Techniques"; pp. 360–72 in *Life Prediction Methodologies and Data for Ceramic Materials*, ASTM STP 1201. Edited by C. R. Brinkman and S. F. Duffy. American Society for Testing and Materials, Philadelphia, PA, 1994.

[8]D. R. Hayhurst, "The Effects of Test Variables on Scatter in High-Temperature Tensile Creep-Rupture Data," *Int. J. Mech. Sci*., **16**, 829–41 (1974).

[9]W. Mendenhall and T. Sincich, *Statistics for Engineering and the Sciences*; Ch. 14.10. Dellen Publishing, San Francisco, CA, 1992.

[10]W. E. Luecke, S. M. Wiederhorn, B. J. Hockey, G. G. Long, and R. F. Krause Jr., "Cavitation Contributes Substantially to Creep in Silicon Nitride," *J. Am. Ceram. Soc*., **78** [8] 2085–96 (1995).

[11]A. A. Wereszczak, T. P. Kirkland, K. Breder, M. K. Ferber, and P. Khandelwal, "High Temperature Dynamic Fatigue Performance of a Hot Isostatically Pressed Silicon Nitride," *Mater. Sci. Eng. A*, **191**, 257–66 (1995).

[12]A. A. Wereszczak, K. Breder, and M. K. Ferber, "Role of Oxidation in the Time-Dependent Failure Behavior of HIPed Silicon Nitride at 1370°C," *J. Am. Ceram. Soc*., **76**, 2919–22 (1993).

[13]A. A. Wereszczak, T. P. Kirkland, and M. K. Ferber, "Difference in Creep Performance of a HIPed Silicon Nitride in Ambient Air and Inert Environments," *Ceram. Eng. Sci. Proc*., **16** [5] 901–909 (1995).

[14]C. J. Gasdaska, "Tensile Creep in an *In Situ* Reinforced Silicon Nitride," *J. Am. Ceram. Soc*., **77**, 2408–18 (1994).

[15]"Standard Test Method for Elevated Temperature Tensile Creep Strain, Creep Strain Rate and Creep Time-to-Failure for Advanced Monolithic Ceramics," C1291-95, 1995 Annual Book of ASTM Standards, Vol 15.01. American Society for Testing and Materials, Philadelphia, PA, 1995. □