# Reliability of the Nasopharyngeal Fiberscope (NPF) for Assessing Velopharyngeal Function: Analysis by Judgment

MICHAEL P. KARNELL, Ph.D. KAORU IBUKI, D.D.S., Ph.D. HUGHLETT L. MORRIS, Ph.D. DUANE R. VAN DEMARK, Ph.D. Iowa City, IA 52242

The reliability and validity of data about velopharyngeal function obtained with the nasopharyngeal fiberscope was assessed in normal subjects. The experimental design included data reduction procedures that are likely to have clinical utility (clinical ratings). The results indicated that relative velar movement and size of the velopharyngeal port may be reliably and validly estimated using the procedures. However NPF estimates of lateral pharyngeal wall movement were not reliable. Finally, the data indicated that velar movement and size of V-P port were consistent within subjects and tasks across data collection sessions. Data about consistency of lateral wall movement across sessions was inconclusive, however. Additional research involving similar procedures with subjects who have morphologic deficits is indicated.

A previous report (Ibuki, Karnell, and Morris, 1983) described the nasopharyngeal fiberscope (NPF), its stability of placement for assessing velopharyngeal function, and some information about reliability and validity of NPF findings when they are analyzed by measurement. Those studies indicated that, with the criterion of acceptability of the NPF view that all boundaries of the velopharyngeal mechanism are visible in a single NPF view, the fiberscope can be placed on repeated occasions in a reliable fashion. The data also indicated that the reliability of a trace-and-measure method of analysis is satifactorily high, except for determination of left lateral pharyngeal wall movement. This exception may be an artifact of our practice of inserting the fiberscope in the right nostril (into the right nasal passageway).

In addition the previous findings indicate that, using measurements from lateral xray films (cinefluoroscopy) as the outside criterion, NPF measurements of velopharyngeal port size and velar movement can be considered valid.

The trace-and-measure method of analysis is not satisfactorily practical in most clinical settings, since it requires preparation of the NPF view in the form of photographic transparency and subsequent projection of the transparency prior to tracing and measurement. All this is time-consuming and constitutes a real limitation to clinical usefulness. It seems reasonable to assume that, to be clinically useful, the NPF procedure must be standardized such that reliable and valid judgments about velopharyngeal function can be made during an examination that is sufficiently brief to be time and cost effective. The purpose of this study was to take an initial step toward determining the efficacy of such procedures by estimating the reliability and va-

The authors are affiliated with the Department of Otolaryngology—Head and Neck Surgery, The University of Iowa, Iowa City, IA 52242. Address Editorial Correspondence to Dr. Hughlett L. Morris.

lidity of clinical ratings applied to the analysis of still NPF photographs of the velopharynx.

#### Methods

**PROCEDURE.** The equipment, data collection, procedures, subjects and films (NPF and lateral cinefluoroscopic) used in this study are the same as those described in our previous report (Ibuki et al., 1983). Briefly, the Olympus nasopharyngeal fiberscope (NPF Type S3), with cold light supply, was coupled to a 35 mm still camera with a mounting adapter, to obtain photographs of the NPF views. The fiberscope was inserted through the right nostril following application of a spray surface anesthesia (4% Lidocaine solution) to the nasal cavity. The fiberscope was positioned to provide, in a single NPF view, all boundaries of the velopharyngeal port. After the fiberscope was properly positioned it was stabilized by a clamp just anterior to the patient's nostril.

The subjects were also prepared for lateral cinefluoroscopy filming which was performed simultaneously with the NPF films. Obviously other radiographic views, such as multi-view procedures, would be more useful than lateral views for the purposes of this study, but were not available. Phillips equipment was used with a fluoroscopic image intensifier. The subjects were positioned for mid-sagittal films including head stabilization with a conventional cephalostat. Appropriate shields were placed for protection of eyes, thyroid, and gonads. The microphone for recording the speech signal was positioned such that the shutter click from the NPF camera was also recorded on the sound track of the cine film. Identification of the cinefluoroscopic film frames that were exposed simultaneously with the NPF films was then possible by locating the sound of the shutter click on the sound track. The cine films exposed while the NPF camera shutter was open were used in the analysis.

Subjects were four adults (three females and one male) who had no evidence or history of speech production disorders or congenital or acquired orofacial or nuerologic anomalies. The decision to use normal subjects was based on the assumption that normal velopharyngeal structures would be less variable than the velopharyngeal structures of subjects with morphologic abnormalities. Since our goal was solely to estimate reliability and validity of NPF assessment **procedures** that are likely to be conducive to clinical settings, evaluation of the procedures with clinical patients was left for subsequent investigations.

Subjects performed the following tasks; quite nasal inspiration, /n/, /i/, /a/, /u/,/s/, /o/, and soft blowing, in that order. They were instructed to inspire between each task in order to allow the velopharyngeal mechanism to return to rest position. The subjects practiced the experimental tasks prior to filming. During filming they were instructed by cue cards.

All subjects were filmed by simultaneous NPF and cinefluoroscopic procedures on the first data collection day. On the same day, the NPF procedures alone were repeated. On a second day (at least one week later) simultaneous NPF and cine films were again taken and, again, NPF films alone were repeated. On a third day (at least one week later), NPF films alone were obtained for two subjects. These various data collection combinations permitted examination of reliability and validity in a manner that minimized radiation exposure to the subjects. Estimated dosage of x-ray by our standard methods of computation was approximately 1.2 rad per subject.

DATA REDUCTION AND ANALYSIS OF NPF FILMS. In order to help the reader visualize the rating system developed for this study, the velopharyngeal mechanism is illustrated schematically as a rectangle (Figure 1) with the four sides representing the velum, right and left pharyngeal walls, and the posterior pharyngeal wall. In the NPF view, the velum was at the bottom, and the other structures were located accordingly. A four point rating scale was employed in order to estimate velar displacement, lateral wall displacement, and velopharyngeal port size.

Rating scales for velar (NVM) and lateral wall (NLW) displacement ranged from -1to +2. A value of zero (0) was assigned for velar and lateral wall positions during quiet



FIGURE 1. Photographic and schematic representation of the velopharynx, as viewed from above, illustrating the rating scales used in this project.

inspiration. A value of plus one (+1) was assigned for moderate displacement toward closure and plus two (+2) was assigned for marked displacement toward closure. A value of negative one (-1) was assigned for displacements away from closure.

The scale for velopharyngeal port size (NVP) ranged from zero to plus three. Zero (0) was assigned when there was no visible opening. Plus one (+1), plus two (+2), and plus three (+3) were assigned for slight, moderate, and marked openings, respectively.

All ratings were assigned relative to the position of the structures during quiet inspiration. NPF films for quiet inspiration were given a mandatory assignment of NVM = 0, NLW = 0, and NVP = +2. All ratings were made relative to these (Figure 2). Investigations involving patients with pathological velopharyngeal mechanisms will need to consider posterior pharyngeal wall displacement. This measure was not included in the analysis of the normal subjects included in this investigation. In addition, for the purposes of this study, we assumed symmetry of lateral pharyngeal wall movement (Iglesias, Kuehn, and Morris, 1980).

TRAINING THE JUDGES. The four judges used in this investigation represented a range of familiarity with NPF procedures and findings. Two were speech pathologists who are highly experienced in cleft palate diagnosis and treatment but who had only limited experience with velopharyngeal examination using the fiberscope prior to this study. The other two, an oral surgeon and a Ph.D. student in speech pathology, had extensive experience with the fiberscope



FIGURE 2. Examples of ratings assigned to various velopharyngeal configurations as viewed by NPF.

and endoscopy prior to the study. The latter two judges also performed the cine tracings and measurement (in this capacity they are identified as tracers in subsequent text).

Immediately prior to the rating session, all four judges viewed a series of nine projected color transparencies (slides) that were developed from one subject's NPF photos. The NPF slides for the subject who most clearly displayed lateral pharyngeal wall movement (as judged by the first author) were chosen for the training procedures. The view of the velopharyngeal mechanism during quiet inspiration for that subject was presented first and declared baseline for that subject. This baseline slide remained on display, side-by-side, as the other NPF slides for the training set were displayed. This method of presentation was then used during all subsequent experimental rating procedures.

During the training session, ratings were given verbally by each judge, the initial rating rotating from judge to judge throughout the series. Such rotation of the initial rating was intended to eliminate the biasing influence likely to occur if the same judge made the initial rating on all slides. Any discrepancies in the ratings between the judges were discussed and debated until at least three of the four judges were in agreement for each slide.

NPF RATING PROCEDURES. Immediately after the training activity, the first set of experimental NPF slides was presented in random order (a "set" is comprised of all NPF films for a single subject, during a single session). The procedure for each experimental set was the same as that for the training set except that each judge recorded ratings on a previously devised rating form, without consultation with the other judges. Each set of experimental NPF slides contained three duplicate slides which were used to estimate intrajudge reliability.

DATA REDUCTION AND ANALYSIS OF CI-NEFLUOROSCOPY FILM. The procedure for cine film analysis was similar to that described by Moll (1960) and Seaver (1978). Cine film frames that corresponded with the NPF films were identified by the evidence of the NPF camera shutter click on



FIGURE 3. Illustration of the templates and measurements for lateral cinefluoros-copy.

the sound track. Templates for the rest position were prepared, and distances were measured from rest position for maximum velar excursion (CVM), velar height (CVH), and minimum velopharyngeal port size (CVP) (Figure 3).

#### Results

Due to poor film quality, cine films for subjects one taken during day 2 and NPF films for subjects two and three taken during day 2 were excluded from the analysis. A summary of the data sets collected is presented in Table 1.

RELIABILITY OF NPF RATINGS. Pearson product-moment correlations were computed to determine how the ratings of each judge agreed with those of each of the other three judges (interjudge reliability). Those correlation coefficients, presented in Table 2, indicated relatively high agree-

 TABLE 1. Summary of Collected Data Sets for

 Each of the Four Subjects

	S #1	S #2	S #3	S #4
Day 1	NPF	NPF	NPF	NPF
	+	+	+	+
Set 1	Cine	Cine	Cine	Cine
Day 1	NPF	NPF	NPF	NPF
Set 2	Only	Only	Only	Only
Day 2	NPF	NPF*	NPF*	NPF
	+	+	+	+
Set 3	Cine*	Cine	Cine	Cine
Day 3	NPF			NPF
Set 4	Only			Only

\* Quality insufficient for analysis.

ment for the ratings of velar movement (.74 to .92) and for velopharyngeal port size (.76 to .92), but indicated less agree-

204 Cleft Palate Journal, July 1983, Vol. 20 No. 3

ment for lateral wall movement (.32 to .66).

Coefficients were computed also to evaluate intrajudge reliability (repeated judgments about the same NPF view). These coefficients were based on data for each parameter, all judges combined, and were as follows: NVM, .98; NLW, .94; and NVP, .97. The relatively high coefficient for lateral wall movement indicated high consistency of rating within judges in spite of the apparent lack of consistency on this parameter between judges.

RELIABILITY OF CINE MEASUREMENTS. As described previously, in this study, each cine frame was traced by each of two tracers independently. Measurements taken from the two sets of tracings were used to estimate the reliability of these measurements. The correlation coefficients computed for each variable between the two tracers were .88 for velar movement (CVM), .86 for size of velopharyngeal

TABLE 2. Correlation Coefficients Estimating Interjudge Reliability for NPF Ratings (n = 109).

NPF Raters	NVM	NLW	NVP
1 vs 2	.82	.52	.82
1 vs 3	.88	.32	.85
1 vs 4	.92	.34	.91
2 vs 3	.74	.32	.76
2 vs 4	.80	.39	.77
3 vs 4	.91	.66	.92

port (CVP), and .71 for velar height (CVH). Intrajudge reliability was not estimated.

VALIDITY OF NPF FINDINGS. The validity of the velar displacement and velopharyngeal port size ratings as assessed by NPF was estimated by comparing these ratings with the measurements of velar movement and velopharyngeal port size from the lateral cine tracings (Table 3). Data collected on day one were employed for this analysis, for each comparison by pairing the cine measurement made by each of the two tracers with NPF ratings obtained from each of the four judges. High correlations (defined arbitrarily at .75 or higher) between these two methods of evaluating velopharyngeal function were considered necessary in order to support the validity of the NPF methods.

Velar Movement (NVM vs CVM). The correlation between velar movement as assessed with NPF rating scale and velar movement as assessed with the cine measurements was acceptably high (.82 to .95). Assessment of velar movement using the nasopharyngeal fiberscope may be considered valid on the basis of this high level of agreement between NPF ratings and cine measurements.

Velopharyngeal Port Size (NVP vs CVP). The validity of the NPF technique for assessing relative size of the velopharyngeal

TABLE 3. Correlation Coefficients Estimating Validity of NPF Ratings Based on Comparisons Between NPF Judges (1, 2, 3, 4) and Cine Tracers (1', 2') (n = 32).

5 0 (	,							
	NPF Ratings vs. Cine Measurements (by judges 1, 2, 3 and 4 and tracers 1'2')							
	1-1'	1-2'	2-1'	2-2'	3-1'	3-2'	4-1'	4-2'
NVM vs CVM	.90	.95	.91	.89	.84	.82	.92	.94
NVP vs CVP	.69	.73	.71	.73	.77	.84	.80	.87
NVM vs CVH	.91	.91	.89	.88	.86	.85	.93	.93
NVM vs CVP	91	94	77	83	89	86	87	92
NVP vs CVM	83	84	83	79	91	92	92	93

port does not appear to be as strongly supported as was that for velar movement. Four of the eight comparisons resulted in correlation coefficients which fell below the .75 acceptance level, and only three of the eight were above .80.

Velar Movement vs Velar Height (NVM vs CVH). NPF validity is supported by the high correlation between velar movement as assessed by the NPF procedures and velar height as measured from the cine tracings. The correlation coefficients for this comparison ranged from .85 to .93.

The correlation between these two variables emphasizes that the vertical component of velar trajectory (velar height), which cannot be viewed directly using the type of NPF view obtained in this investigation, is interrelated with anterior-to-posterior component of velar movement, which can be directly viewed with the NPF. It seems reasonable to conclude that a high rating of velar displacement as assessed with the NPF is associated with increased velar height in the subjects used in this investigation.

Velar Movement vs Velopharyngeal Port Size (NVM vs CVP). The correlation coefficients computed for comparisons between NVM and CVP indicate a high negative relationship between these variables as would be expected. The correlations produced values that ranged from -.77 to -.94. These results may also be interpreted as supporting the validity of the assessment of velar movement by NPF.

Velopharyngeal Port Size vs Velar Movement (NVP vs CVM). Given the high negative correlation between NVM and CVP, a high negative relationship between NVP and CVM would also be expected. As indicated in Table 3, such a relationship did exist between these variables. The correlation coefficients ranged from -.79 to -.93, further supporting the validity of NPF assessment.

NPF Validity Based on Pooled Data. The computations for the previous consideration of NPF validity were based on 32 samples of NPF ratings and cine measurements for each comparison. In an effort to increase the power of the correlations, the analysis procedures were modified so that the number of samples for each comparison was increased. The ratings of NPF judges three and four were combined and compared with the combined cine measurements from tracer one and tracer two. These pairings were selected because raters three and four were most consistently in high agreement with each other in the previous analyses (Table 2). By pooling the data, a total of 64 samples was included in each comparison. The description of the results of these comparisons follows.

The correlations between the NPF and cine variables in the pooled data analysis (Table 4) appear very similar to those reported for the smaller sample size (Table 3). All of the coefficients were above the .75 acceptance level except those involving NLW (the NPF rating of lateral wall movement). An interesting observation is that the negative relationship between NVP and CVM appears to be stronger than the positive relationship between NVM and CVM. Assessment of size of velopharyngeal port by NPF may, therefore, be a more accurate indicator of velar movement than direct assessment of velar movement by NPF. This possibility seems reasonable because of the sharp contrast between the velopharyngeal opening and surrounding tissues. Changes in velar position may, therefore, be more easily identified and reported as they are reflected in velopharyngeal port size modifications than in actual velar displacement.

The coefficients reported in Tables 3 and 4 support the validity of NPF findings analyzed in this manner for evaluating relative displacement and size of velopharyngeal port. Validity of lateral wall displacement could not be analyzed using the experimental design employed in this project because lateral cines do not display lateral wall movement. While lateral wall movement may be expected to be highly correlated with velar movement of velopharyn-

TABLE 4. Validity of NPF Based on Comparisons Between Pooled NPF and Cine Data (n = 64).

		· · · · · · · · · · · · · · · · · · ·	/-
	CVM	CVH	CVP
NVM	.89	.89	85
NLW	.43	.27	34
NVP	92	88	.82

206 Cleft Palate Journal, July 1983, Vol. 20 No. 3

geal port size, the data in the present study do not examine this possibility.

CONSISTENCY OF VELOPHARYNGEAL FUNCTION. Given that the NPF rating system is sufficiently reliable and valid, questions addressing the functional consistency of the validity that affect velopharyngeal closure may be considered. Correlation between NPF ratings of repeated samples of velopharyngeal closure performed on the same day are presented in Table 5, and on different days in Table 6.

Within Data Collection Days. The correlations between data sets 1 and 2, collected on day 1, indicate that the pattern of closure for each of the subjects was generally similar (Table 5). Velar movement and size of velopharyngeal port, in particular, demonstrated consistency between sets. While the correlation between sets for lateral wall movement was relatively low compared to NVM and NVP, the difference may have been due to the lower degree of reliability associated with NLW mentioned above. To assume that lateral wall movement occurs in widely varying degrees based on these results is, therefore, inadvisable.

Between Data Collection Days. NPF recordings were repeated for two subjects on each of three separate days (day 1, set 1, day 2, set 3; day 3, set 4) (see Table 1 for the

TABLE 5. Consistency of NPF Measures of Velopharyngeal Function Within Data Collection Days, Set One and Set Two (n = 92).

Set One		Set Two	
	NVM	NLW	NVP
NVM	.79		
NLW		.59	
NVP			.89

summary of data sets collected). Table 6 displays the correlation coefficients between days for these subjects. As mentioned above, NPF slides that were taken for the other two subjects on one of the two trial days were not usable due to excessive glare and so the data for these two subjects were omitted from this analysis.

The between-day correlation coefficients generally agree with the within day correlations. NVM and NVP appear to be high while NLW is, by comparison, quite low. This apparent trend was consistent for all three comparisons; i.e., day 1 vs day 2, day 1 vs day 3, and day 2 vs day 3.

All obtained coefficients indicate relatively high agreement for velar movement, and velopharyngeal port size, but not for lateral wall movement. As indicated by our previous findings (Ibuki, et al., (1983), repeated placement of the fiberscope gives findings about the velopharyngeal mechanism that are comparable.

Some data are available about consistency of velopharyngeal activity from the cine films. The between-days comparison based on the cine data indicates consistency of performance of each component of velopharyngeal function measured by this method. CVH and CVP both received correlation coefficients of .88, while CVM received a correlation coefficient of .85.

# Discussion

In general the findings from this study indicate that, under certain conditions and with certain restrictions, views of the velopharynx obtained with the nasopharyngeal fiberscope are reliable and valid when analyzed by clinical judgment.

One condition is that the NPF view must be such that all boundaries of the velophar-

TABLE 6. Consistency of NPF Measures of Velopharyngeal Function Between Data Collection Days 1, 2, and 3 (n = 64).

Variable day	NVM 1	NLW 1	NVP 1	NVM 2	NLW 2	NVP 2
NVM 2	.80		•			
NLW 2		.49				
NVP 2			.81			
NVM 3	.89			.76		
NLW 3		.54			.27	
NVP 3			.82			.79

ynx are visible in a single view. This requirement may be an important limitation in the clinical setting: in our subject selection we rejected four of eight subjects screened for NPF films because the velopharynx was so large that this condition could not be met. Another condition, described in our earlier paper (Ibuki et al, in press), is that the subject and fiberscope be both positioned and stabilized as described here or in a similar manner.

A final condition is that the examiner who used the nasopharyngeal fiberscope must have prior training in evaluating the NPF views. We used a specific training session for that purpose. In addition, two of our four examiners had considerable experience with the fiberscope while two had not. Some differences in reliability between the less experienced and more experienced groups were observed but they were not consistent.

There are limitations to these data that must be considered in current use of the NPF and in planning future research with it. We used only normal adult subjects in order to limit intersubjects variability. Therefore, inferences from findings with these subjects to other populations must be made with caution.

Some of our NPF films were not usable because they showed a high degree of glare. Presumably, that glare was not apparent in the view through the fiberscope eyepiece before the NPF film was taken or else adjustments would have been made. The problem of glare should be considered further in future investigations.

Our initial interest in the fiberscope was based in part on the likelihood that NPF views would be informative about lateral wall movement during speech and related activities, and in a relatively non-invasive manner. The data reported here indicate that these clinical ratings of NPF views are not satisfactorily reliable for that purpose. We assumed symmetry of lateral wall movement and asked for one rating covering both lateral walls combined. Perhaps reliability would be higher if we had asked judges to focus on only one wall at a time. The data from our previous study (Ibuki et al., 1983), in which NPF films are traced

and measured, indicate high reliability for the right lateral wall but not the left lateral wall. We speculated in that paper that the difference might be due to the fact that we always inserted the fiberscope into the right nasal passageway.

# Conclusions

The data obtained in this investigation indicate that findings with the nasopharyngeal fiberscope about velopharyngeal function are informative, and that the fiberscope procedure, as described here, can be considered to be a useful procedure in assessment of the velopharyngeal mechanism and function. Data about reliability and validity of NPF findings with clinical patients are needed, and we urge that such research be conducted in the immediate future.

Until information is available about the effect of unstabilized NPF examination technique on obtained results, we urge also that our procedure for positioning and stabilizing the fiberscope or some close modification of it, be used.

We urge also that examiners, regardless of discipline, spend considerable time becoming familiar with the nasopharyngeal fiberscope and NPF views before they begin to rely on their findings. Furthermore, Willis (1981) has recommended, and we agree, that a set of NPF slides be kept handy for occasional judge re-calibration in order to maintain internal consistency.

Finally, we remind potential users of nasopharyngeal fiberscope that like several other physiologic and instrumental techniques, the fiberscope yields information about structures and structural movement, and not the behavioral basis for the movement. For example, a patient who is capable of producing an oral plosive but typically uses a nasalized plosive will show velopharyngeal closure on the first but not on the second by NPF or any other physiologic measure. Which one is valid? They both are, but both must be interpreted within the context of the patient's repertoire of speech production. Like other measures of velopharyngeal function, the nasopharyngeal fiberscope, though it yields interesting and useful information, must be used in coordination with other procedures that provide information about the behavioral aspects of speech production. The importance of involvement of the speech pathologist is, therefore, emphasized.

Acknowledgment: The authors thank Mr. H. Morishima and the Olympus Corporation of America, New York for their courtesy in providing the nasopharyngeal fiberscope for our project.

Thanks are also extended to Dr. Clyde R. Willis of Western Michigan University, Kalamazoo for generously sharing the benefits of his extensive experience in the clinical use of endoscopy.

This study was supported in part by Pub-

lic Health Service Grant DE-0853 from the National Institute of Dental Research.

### References

- IBUKI, K., KARNELL, M. P., and MORRIS, H. L., Reliability of the nasopharyngeal fiberscope (NPF) for assessing velopharyngeal function. *Cleft Palate J.*, 20: 1983
- IGLESIAS, A., KUEHN, D. P., MORRIS, H. L., Simultaneous assessment of pharyngeal wall and velar displacement for selected speech sounds, *J. Speech Hear. Res.*, 429–446, 1980.
- Moll, K. L., Cinefluorographic techniques in speech research, J. Speech Hear. Res., 227–241, 1960.
- SEAVER, E. J., A. cinefluorographic and electromyographic investigation of velar positioning in oral speech, Doctoral Dissertation, University of Iowa, 1978.
- WILLIS, C. R., Personal communication, 1981.