A review of important elements in sedation study methodology

Stephen Wilson, DMD, MA, PhD

The purpose of this article is to offer a historical review of sedation studies with the goal of sensitizing the reader to key issues related to behavioral parameters in the dental operatory.

Overview

Although literature describing the use of pharmacologic agents for sedation exists from before 1950, it will be not reviewed. The majority of articles since then describing sedated children in the dental operatory were published after 1980. The Table displays the majority of sedation articles published in Pediatric Dentistry or the Journal of Dentistry for Children.

Types of behavioral measurements

The types of behavioral scales used in these studies can be categorized broadly as either global (e.g., Frankl Scale) or restricted (e.g., Houpt Scale). Global scales utilize an underlying measure that is discrete, but categorical, such as “definitely negative,” which denotes a child who exhibits disruptive, uncooperative behaviors with crying. The Frankl scale allows observed behavior occurring over a given period of time to be grossly interpreted and recorded into one of four discrete categories that are mutually exclusive.

The advantages of these global scales are their simplicity and purported implication for the clinical context and appreciation by most practitioners for any particular pharmacologic protocol. Some drawbacks are:

1. The increased possibility of “lost” information during key procedures of the sedative visit (e.g., it is improbable within such a scale that a rater can perceive, register, recall, collate, and record subtle or dramatic changes in behavior over time within the configuration of four response categories)
2. “Halo” effects in which the more dominant behavior prevailed in rating outcomes, although many behaviors may have occurred during the rating session
3. A lack of demonstration of standardization (i.e., reliability and validity) of the scale within and across studies, and
4. The application of less powerful statistics (only nonparametric statistics can be used and most clinicians are not familiar with the use of median and modes rather than the mean as measures of central tendency).

Often with these scales, the mean is reported as a measure of central tendency. What does a mean of 1.67 on the Frankl Scale mean? Is that a group of children whose behavior is slightly closer to, but between “definitely negative” and “negative” categories? The scale implies the behavior can only be one or the other as each is a separate category. It’s a useless measure in this circumstance.

Restricted scales may use measures of either discrete (e.g., Houpt Scale) or continuously occurring (e.g., Ohio State University Behavioral Rating Scale [OSUBRS]) specifically defined and limited behaviors. For instance, the Houpt Scale has three major categories (i.e., degrees of movement, wakefulness/sleep, and quiet/crying). The major categories are divided into three or four subcategories, each varying in the degree of expression in a rank-order fashion. For the major category of “wakefulness/sleep,” three subcategories, “fully awake,” “drowsy, disoriented,” and “asleep” range in score from 1 to 3, respectively.

Advantages include:

1. The perception, discrimination, and recording of different types of behaviors varying in degrees of expression within a given recording period
2. Increased likelihood of capturing subtle, but clinically important behaviors
3. A more complete accounting of behaviors (higher informational content per unit of time) including frequency and often, the duration of occurrence of a definite behavior
4. The use of more powerful statistics for data analysis.

There are a few disadvantages, however. One that is significant is the rater’s ability to recall, discriminate, and assign a rank among several choices of subcategories for more than one major category while viewing a sedation. Unless videotaping is used for repeated viewing, accuracy may suffer.

"Review Article"
<table>
<thead>
<tr>
<th>Decade</th>
<th>Author</th>
<th>Age (Years)</th>
<th>No. of Kids</th>
<th>Drug Combo of Drugs</th>
<th>Monitoring</th>
<th>Physiology</th>
<th>Behavior</th>
<th>Reliability</th>
<th>Blinded</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>50s</td>
<td>Buckman</td>
<td>2.5-9</td>
<td>33</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Poor</td>
</tr>
<tr>
<td>60s</td>
<td>Kracke</td>
<td>&lt; 4-15</td>
<td>255</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% Success</td>
</tr>
<tr>
<td>60s</td>
<td>Czarnecki &amp; Binns</td>
<td>1-9</td>
<td>100</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Poor</td>
</tr>
<tr>
<td>60s</td>
<td>Robbins</td>
<td>2-6</td>
<td>48</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>Good-Poor</td>
</tr>
<tr>
<td>60s</td>
<td>Chalmers &amp; Galbreath</td>
<td>3-12</td>
<td>38</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Poor</td>
</tr>
<tr>
<td>60s</td>
<td>Jones</td>
<td>?</td>
<td>100</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cooperation</td>
</tr>
<tr>
<td>60s</td>
<td>Anderson</td>
<td>?</td>
<td>300</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Satisfactory-Poor</td>
</tr>
<tr>
<td>60s</td>
<td>Evans et al</td>
<td>3-8</td>
<td>75</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>Emotions-Manageable</td>
</tr>
<tr>
<td>70s</td>
<td>Kopel</td>
<td>?</td>
<td>44</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Poor</td>
</tr>
<tr>
<td>70s</td>
<td>Tobias et al</td>
<td>2-14</td>
<td>124</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Effective-Poor</td>
</tr>
<tr>
<td>70s</td>
<td>Smith</td>
<td>4-16</td>
<td>14</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Frankl &amp; Interference</td>
</tr>
<tr>
<td>70s</td>
<td>Myers &amp; Shoaf</td>
<td>2-5</td>
<td>26</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Unsatisfactory</td>
</tr>
<tr>
<td>70s</td>
<td>Bragg &amp; Nealon</td>
<td>12-24</td>
<td>14</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70s</td>
<td>Lundgren et al</td>
<td>3-7</td>
<td>33</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Treatment Acceptance</td>
</tr>
<tr>
<td>80s</td>
<td>Barr et al</td>
<td>4-8</td>
<td>21</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Unsatisfactory</td>
</tr>
<tr>
<td>80s</td>
<td>Auil et al</td>
<td>4-8</td>
<td>60</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent-Fair</td>
</tr>
<tr>
<td>80s</td>
<td>Moore et al</td>
<td>2-8</td>
<td>50</td>
<td>√</td>
<td>√</td>
<td>HR, RR, BP</td>
<td>√</td>
<td></td>
<td></td>
<td>Satisfactory-Unsatisfactory</td>
</tr>
<tr>
<td>80s</td>
<td>Mueller et al</td>
<td>2-6</td>
<td>40</td>
<td>√</td>
<td>√</td>
<td>BP, RR, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Wilson et al</td>
<td>2-3</td>
<td>10</td>
<td>√</td>
<td>√</td>
<td>BP, HR, RR, CO2, Ox, EMG</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Doring</td>
<td>2-10</td>
<td>26</td>
<td>√</td>
<td>√</td>
<td>BP, HR, Ox, RR</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Flaitz &amp; Nowak</td>
<td>2-6</td>
<td>12</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Houpt et al</td>
<td>1.5-3.5</td>
<td>21</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sleep, Move, Cry</td>
</tr>
<tr>
<td>80s</td>
<td>Schneider</td>
<td>1-&gt;14</td>
<td>4363</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Nathan &amp; West</td>
<td>1.5-5</td>
<td>135</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Lambert et al</td>
<td>2-5</td>
<td>28</td>
<td>√</td>
<td>√</td>
<td>BP, HR, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Whitehead et al</td>
<td>2-5</td>
<td>24</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Currie et al</td>
<td>2-9</td>
<td>14</td>
<td>√</td>
<td>√</td>
<td>BP, RR, HR, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Iwasaki et al</td>
<td>2-5</td>
<td>10</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Moody et al</td>
<td>2-5</td>
<td>60</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s</td>
<td>Moody et al</td>
<td>2-6</td>
<td>30</td>
<td>√</td>
<td>√</td>
<td>BP, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Badalaty et al</td>
<td>1.5-4</td>
<td>30</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Walbergh et al</td>
<td>1.5-4</td>
<td>18</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Wilson</td>
<td>1.5-3</td>
<td>22</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Wilson et al</td>
<td>2-5</td>
<td>10</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>McKee et al</td>
<td>2-5</td>
<td>60</td>
<td>√</td>
<td>√</td>
<td>BP, HR, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Poorman et al</td>
<td>2-5</td>
<td>40</td>
<td>√</td>
<td>√</td>
<td>HR, RR, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Meyer et al</td>
<td>2-6</td>
<td>40</td>
<td>√</td>
<td>√</td>
<td>HR, RR, Ox</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90s</td>
<td>Hasty et al</td>
<td>2-4</td>
<td>10</td>
<td>√</td>
<td>√</td>
<td>HR, RR, Ox, CO2</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Continued on next page
Of the 57 articles in the table, 46 focus on behavior and, of those, 33 report the measurement of some physiological parameter. At least eight different behavior rating/scoring scales have been used in the 46 articles. The most frequently used scale (13 articles) is a type involving a categorical-range scale of excellent to poor, with the second and third most frequently used being Houpt (eight articles) and the Frankl (five articles) scales. The remainder consist of a miscellany of scales (e.g., North Carolina Behavior Rating Scale).

**Blinding and reliability**

Blinding refers to the process of removing from the rater access to significant information that may bias the outcome of a study. For example, knowledge of the type of drug used in a comparative study might bias an investigator if that drug is a "favorite".

Reliability refers to the process of determining the degree of association when a variable is repeatedly measured by one or more individuals or when the same operationally defined behavior is observed/rated simultaneously between two or more individuals. Often studies use complicated scales and do not indicate that raters were skilled in their use. In most legitimate psychological studies, investigators go to great lengths to calibrate their raters and standardize scales against those of known reliability.

Fewer than half of the 46 articles (22) using behavioral measures utilized the principle of operator and/or rater blinding, and only 11 used a combination of blinding and some form of reliability measure of the behavioral scale used to rate the child's responses. Thus, out of the cumulative total of 7,201 children (only 2,838 if the Schneider article is eliminated), only 11 studies involving 373 (5%) children were rated using a behavioral measure or scale that was judged reliable and the rater blinded. Of these 11 studies, six different scales were used, with Houpt being the most common.

Blinding and reliability measures are important considerations in studies involving behavioral assessments of complicated behaviors such as those found in sedated children. Reliability measures give an indication of ease of application of the measure being evaluated. If a measure tends to be highly reliable, either a great
deal of investment was given the training of the raters or the measure is probably nondiscriminatory and questionable. If the degree of reliability is not established in a study, the readers must decide the value and applicability of the study to their clinical experience.

Blinding eliminates the majority of internal (personal) and external (interpersonal) biases of key personnel in the study. Blinding was accomplished in 21 of the 57 articles in the Table. Study results should only reflect the effects of the variable manipulated; blinding prevents their tainting. Whether acknowledged or not, practitioners manifest biases in many of their philosophies, expectations, and procedures. Without minimizing the influence of bias in studies, separating out effects of key variables becomes moot.

For example, assume an investigator conducted a study looking at the effect of two different agents on the amount of disruptive behavior and did not use any blinding techniques. Further, the investigator believed one of the agents worked most effectively, based on prior experience (but this belief is still unproven). The investigator rates videotaped trials with knowledge of which agent was used. The subtle effects of the "belief" cannot be separated from the judgement process in this case. The bias may insidiously manifest itself as a perception of a decreased amount of crying with the drug of choice or possibly manifest as a cry whose salient effect is perceived as more shrill and thus the non-preferred drug assumed as less effective.

Blinding and reliability procedures are not the only important elements in sedation studies. Other considerations may include the number of patients involved, length of the sedations, patient selection process, control groups/conditions, and procedural effects. Nonetheless, blinding and reliability together are key elements of the standardization of every study that should not be manipulated.

Studies that did not use blinding and reliability measures are not without some pragmatic value. For instance, the description of a new rating scale, the accentuation of a scale’s or a drug’s advantages and disadvantages, the comparison of multiple behavioral scales within a single study, or the suggestion following a trial basis that a drug effect is measurable and applicable are important and useful considerations. However, interpreting unblinded studies becomes difficult in any scientific sense and, importantly, may misguide a practitioner’s expectations and routines during sedation appointments.

Clinical significance of rating scales

Today’s research arena of conscious sedation typically involves the use of behavioral scales containing multiple measurable responses such as degree of quietness or sleep, crying, movement and struggling (e.g., HoupT Scale). Children’s responses are usually recorded for brief periods during key procedures (i.e., injection) and/or on a time-based paradigm (e.g., every 5 min).

Such a paradigm gives the reader a more complete perspective and concept to which clinical experience can be related. In simplistic terms, its like watching a sunset (global scale) versus watching a boxing match (restricted scale). Furthermore, evidence is accruing that the majority of disruptive behaviors occurs early during a sedation and are associated with more painful stimuli (e.g., injections). To lump all behaviors during a sedation into a global category does an injustice to the practitioner seeking to discriminate among regimens or procedures. A time-based or procedural picture of a sedated child’s behavior and physiology has more clinical meaning in the context of anticipating, planning, and performing a sedation than does a global impression of what one thought was observed.

A majority of recent reports has used various operationally defined measures of specific behaviors and has assessed physiological responses during sedation trials. No doubt, this trend reflects:

1. The need for more reliable measures and a generalized assessment of behavioral responses of children (with both contributing relevance to the private practice setting)
2. Advances in videotaping and controlled playback
3. Development of specialized computer software
4. Development of sedation guidelines
5. Electronic monitoring sophisticatedly packaged in modular units.

In contrast, many earlier studies consistently used more global scales (e.g., Frankl Scale or categorical ranges from excellent to poor) and less operationally defined behaviors applied as indices of the child’s behavior over the entire sedation visit, although current studies occasionally use such indices. It is critical that one be aware of the rating scale used in sedation studies, especially if the goal is to formulate some clinical sense of pertinence and comparison within a procedure that’s as much an art as a science. Not only do the intrinsic measures differ from scale to scale (e.g., the North Carolina Behavior Rating Scale [NCBRS] reflects only disruptive types of behaviors and its use in any given trial would eliminate the accounting of nondisruptive or cooperative behaviors), but the scoring process, or how an examiner accounts for the behaviors rated, varies considerably between scales.

For instance, the Ohio State University Behavioral Rating Scale (OSUBRS), a modification of the NCBRS, uses a computer program to capture mutually exclusive classifications of behavior (e.g., crying versus quiet) recorded continuously over time. Typically, the rater views a videotape while sitting at a computer and pushes specified letters (e.g., a “c” for “crying”) on the keyboard that trip the counting and timing of a designated code representing a behavior. Each behavior is mutually exclusive from the others, that is, a child cannot be quiet and nonmoving while simultaneously cry-
ing and thrashing about in the dental chair. The computer tallies the number of times and duration of a given code.

Typically, a summary printout using the OSUBRS may show a child who displayed 73 sec of quiet behavior before crying for 15 sec, then followed by 35 sec of struggling. The outcome of the OSUBRS is the percentage of a given behavior rated over the duration of the rating session. In the example above, quiet, crying, and struggling behavior represented 59, 12, and 29%, respectively of the total duration of time rated.

This is an important concept to appreciate. When one sedation trial varies in length of time from another and this occurs more frequently than not, this type of scale or analysis is useful. It gives the reality of a mixture of behaviors, each of which varies in the frequency and duration of occurrence. It also standardizes the length of each sedation trial, making statistical and clinical analysis more meaningful.

A disadvantage of this scale is that some practice time on the part of the rater(s) is necessary both to familiarize the rater with the task and to develop a high degree of reliability. Statistical analysis with this type of scale permits the valid use of parametric tests (e.g., ANOVA). The “mean” duration of crying has some meaningful implications for a practitioner, but the practitioner must be cognizant of the effect of variance (i.e., the contribution of extremes) on the mean. Variance is a measure of data distribution not all of which are normal (bell-shaped). As a practical example, in a sample of 10 children only two or three of whom cry for a protracted period of time compared to the remaining children, the mean may be spuriously affected.

In comparison, the same behavior when scored with the Houpt Scale requires a rater to view or recall either mentally or via videotape review the child’s behavior three times with each review devoted to the wakefulness/sleep, crying, and movement score, respectively. Repeated videotape viewing is preferred, otherwise one risks recalling from memory a score for each of the categories when the behavior is viewed only once. For example, a brief but explosive behavior sequence may color the reviewer’s rating, but be only a brief part of a sedation session. Repeated viewing or recall mechanisms on the other hand, may risk the development of bias affecting scores across the categorical measures used.

The outcome of the Houpt Scale is a sum of classifications for each rating period, so it represents a miniaturized, flexible, and more descriptive Frankl scale. The measurement is discrete and permits some conceptualized, meaningful ranking (e.g., a violent, interrupting movement is more disconcerting to the operator than a controlled movement that does not interfere with treatment), yet the difference in magnitude between these behaviors in any clinical sense may be either difficult or obvious to assess, depending on the situation.

The statistical analysis of such a scale requires a nonparametric test (e.g., Mann-Whitney U). Thus the outcome measure, for example, the median (which is a measure of central tendency in nonparametric, categorically ranked measures) often has less meaningful implications for the practitioner. Also, the median is not affected by the extremes in a sample as is the mean. Nonetheless, the categories used in the Houpt Scale are easily recognized and appreciated by either the neophyte or seasoned rater/operator. This characteristic makes it a highly desirable scale to make comparisons across studies.

In summary and comparison, the outcome for the OSUBRS versus the Houpt scale (and these are only taken as examples) is a mean duration during which a behavior is exhibited (OSUBRS) compared to the assigning of a representative category of behavior either during a procedure or throughout a rating session, respectively (Houpt). For the same time frame one is frequency- and duration-related, showing fluctuations in behaviors (OSUBRS), while the other is a group description of rank-classified behaviors (Houpt), like a “snapshot gestalt” of a scene. One is dynamic the other static.

Clinical relevance and recommendations

Even though behavioral studies may be the most time consuming and technically demanding to conduct, it is necessary to view the behavioral scoring paradigm as a critical issue deserving attention and control. The essence of the scientific method, its control, and limitation of interpretation, set the standard against which any manipulation of drugs can have any clinical association and significance.

What significance does a body of clinical knowledge have when derived from methodologies that vary in measure and interpretation depending on one’s experience or need for prediction of child behavior? Quite frankly, it is difficult, if not impossible, to make meaningful comparisons among sedation study outcomes with this variance in scales and protocols. The scales differ in type of response measurement, processing, and implication. Protocols that are not standardized decrease the relevance for the practice situation. On the other hand, it is not always feasible to use only one scale across studies because the outcomes desired will vary depending on the purpose of the studies. As proffered solutions, one may consider using a single scale used on a standardizing basis as a preference for a specific outcome or purpose (e.g., drug dose-response studies). Alternatively, more than one scale can be used in any given study to help demonstrate the validity of the study’s purpose while providing diversity for interpretive purposes in attaining the study’s goal.

Summary

The table demonstrates that there are few studies of behavioral indices mediated by a sedative(s) involving
such sound scientific principles as blinding, reliability and validity, and control groups. Indeed, as is reflected in this report, it is not common to have measured physiological parameters along with the behavioral indices. Nonetheless, the reader should be aware that these clinical studies are difficult to conduct under the best of conditions. Extensive planning, piloting, instruments, and time are necessary to render even simple conclusions in today's studies.

The need to standardize sedation protocols within and across studies is great if the profession ever hopes to develop a strong scientific basis for sedation. Only then can studies of drugs, their dosages, and their effects on children during dental procedures be compared and contrasted with any confidence. Behavioral scales need to be assessed and compared repeatedly and independently within and among studies so that an appreciation of the influence of a drug on one scale will have some meaningful translation to another scale. For instance, how does the Frankl scale compare to the Houpt scale, if at all? Should both or some facsimile of each be included in every study? What type of information is derived from the use of a given scale and does that information impact on other factors (e.g., physiological parameters or number of quadrants completed)? Should there be a priority for the use of one scale or another depending on the purpose of the study? Blinding and the establishment of reliability of behavioral measures should become as second-natured as giving local anesthesia.

Dr. Wilson is associate professor and director, postgraduate pediatric dentistry program and research, The Ohio State University, Columbus, Ohio.  


Send us your manuscript on disk...

To expedite publication of manuscripts in Pediatric Dentistry and to reduce the chance of introducing errors, we now require that all submissions include a diskette. Mark clearly on the disk the type of computer (Macintosh, IBM, etc.) and the word processing software (Microsoft Word, WordPerfect, etc.) you used. Please continue to send four copies of the text, and four originals of all photographs and figures.

Thanks for your cooperation.