*North American Journal of Fisheries Management* 15:49–53, 1995


# Observer Variability in Classifying Habitat Types in Stream Surveys

BRETT B. ROPER AND DENNIS L. SCARNECCHIA
*Department of Fish and Wildlife Resources, University of Idaho, Moscow, Idaho 83843, USA*

*Abstract.*—We report on the ability of trained observers to independently classify habitat units within stream reaches into primary (pools, riffles, and glides) and secondary (types of pools and types of riffles) habitat types. Differences among observers in classifying habitat types increased with the number of habitat types and decreased with level of observer training. Observer variability also seemed to be affected by reach-specific physical attributes, such as gradient and the amount of wood in the stream channel. Attempts to classify stream habitats will be more consistent and useful if observers receive sufficient uniform training and are required to distinguish between fewer habitat types.

Physical habitat evaluations are often an important tool for assessing the effects of human activities on a stream and its biota (Heifetz et al. 1986; Hicks 1989). Streams consist of continuous reaches of flowing water naturally divided into distinct habitat units (Leopold et al. 1964), such as pools and riffles and glides with varying depths, velocities, and substrate types. The distribution and frequency of specific habitat units within a basin reflect basin geomorphology (Leopold et al. 1964; Knighton 1984; Montgomery and Buffington 1993) and influence the biotic community (Dambacher 1991; Bisson et al. 1982). Therefore, stream evaluations are often stratified according to habitat types identified visually by trained independent observers (Hankin 1984; Hankin and Reeves 1988; Hawkins et al. 1993).

For stratification by habitat type to be useful, independent observers must be able to classify habitat units objectively and consistently (Southwood 1980). Independent observers must also be able to estimate dimensions of the habitat units once the units have been identified (Hankin and Reeves 1988). We report on three trials that evaluate the ability of trained observers to independently classify habitat units into primary (pools, riffles, and glides) and secondary (types of pools and types of riffles) habitat types, and to visually estimate stream channel dimensions.

## Methods

Classification trials were conducted on a single reach on each of three streams within the South Umpqua River basin of southwestern Oregon. Trial 1 was conducted in 1992, and trials 2 and 3 in 1993. Two of the three trials were in basins in which timber had been harvested from more than 25% of the area, and the third trial was in an undeveloped basin (Table 1). All three trials were completed during summer when streams were at base flow.

*Trial 1: Dumont Creek.*—Eight trained observers classified habitat units into types on the same reach of Dumont Creek. All observers had received sufficient training so that the U.S. Forest Service (USFS) previously and subsequently used them to conduct stream surveys without additional training. Five of the eight observers had been trained in a 3-d USFS short course, of which 1 d was spent in the field classifying habitat units into types. We provided no additional training.

The observers classified the stream reach into primary habitat types (USFS 1992) and visually estimated the length and average width of each habitat unit, sequentially in an upstream direction on the same day. Total length and area of the trial reach were estimated by the methods described by Hankin and Reeves (1988). Observers were also instructed to further classify pools into secondary habitat types: trench, plunge, backwater, dam, and lateral scour as defined by Bisson et al. (1982). All observers were initially given an instruction sheet that defined primary and secondary habitat types and were permitted to refer to these instructions during the trial. A habitat unit was considered to be discrete only if its length exceeded its average width; otherwise, observers were instructed to include that unit's surface area in an adjacent habitat unit.

After the observers finished, actual measurements of the habitat units were made by the senior author. Length was measured through the center of the stream channel and habitat widths were measured every 2–10 m, depending on the stream channel complexity and the length of a particular habitat unit. Visual estimates were compared to

TABLE 1.—Characteristics of the watersheds in which the three habitat classification trials were conducted. Large woody debris (LWD) are pieces of wood within the active channel greater than 60 cm in diameter and 8 m in length.

| Variable | Trial 1 | Trial 2 | Trial 3 |
|---|---|---|---|
| Creek | Dumont | Slick | Lonewoman |
| Stream order | 4 | 3 | 3 |
| Stream width (m) | 6.4 | 5.8 | 4.9 |
| Watershed area (ha) | 8,099 | 4,587 | 1,966 |
| Percent logged | 28 | 29 | 3 |
| Number of LWD/km | 4 | 6 | 24 |
| Gradient (%) | 2.3 | 2.5 | 4.0 |

actual measurements and correction factors and confidence intervals of visual estimates were calculated with equations 3 and 4 from Hankin and Reeves (1988).

We then compared how each observer had classified the reach into primary and secondary habitat units. For each habitat unit that at least five of the eight observers classified as a pool, we compared how the observers differed in their classification of that habitat unit into different secondary habitat types. Because there was no "correct" interpretation, we investigated variability (range and coefficient of variation) in the observers' classifications, not the mean values.

*Trials 2 and 3: Slick Creek and Lonewoman Creek.*—Because few habitat units (mean, 19) were identified in Dumont Creek in 1992, additional trials were conducted by six observers on two other streams in 1993. The two streams were chosen to reflect different levels of land management activities (Table 1). All observers had been professionally involved in habitat classification surveys periodically for at least 2 years. Consistency among observers was enhanced by providing them with uniform training at a 3-d USFS training session, in which the observers reviewed habitat survey and classification methods. Also, the group spent 2 d in the field classifying stream habitats.

In each trial, 75 habitat units were identified (25 each by three of the observers in each stream) for classification by habitat type. Each of the six observers classified the 150 habitat units into the three primary habitat types and five pool types identified in Trial 1. Additionally, riffles were classified as low-gradient, high-gradient, or cascading (Bisson et al. 1982). As in trial 1, we assessed agreement among observers by investigating the variability in classifications.

Complete agreement among observers in classifying a habitat unit's primary or secondary habitat type was used to indicate consistency. Chi-square goodness-of-fit tests were used to determine if complete agreement among observers in classifying primary and secondary habitat types were equally likely. Frequencies of habitat types within a reach were estimated by assigning to each habitat unit the type designated by the majority of the observers.

TABLE 2.—Number of habitat units identified by each observer, corrected visual estimate of total reach area, and percentage of the area in each of the three primary habitat types in trial 1. The actual measured area was 2,717 $m^2$. Correction factors and confidence intervals (CI) were determined with equations from Hankin and Reeves (1988).

| Observation | Number of units | Area correction factor | Corrected area ± 95% CI | Percentage of area classified as: Pool | Riffle | Glide |
|---|---|---|---|---|---|---|
| 1 | 19 | 1.41 | 2,820 ± 566 | 61 | 39 | 0 |
| 2 | 16 | 0.96 | 2,576 ± 402 | 52 | 38 | 10 |
| 3 | 17 | 1.12 | 2,885 ± 347 | 34 | 44 | 22 |
| 4 | 21 | 2.23 | 2,681 ± 348 | 63 | 35 | 2 |
| 5 | 15 | 0.83 | 2,602 ± 361 | 66 | 31 | 3 |
| 6 | 21 | 0.83 | 2,825 ± 191 | 57 | 42 | 1 |
| 7 | 18 | 1.07 | 2,407 ± 678 | 52 | 35 | 13 |
| 8 | 20 | 0.88 | 2,686 ± 694 | 60 | 39 | 1 |

## Results

### *Trial 1: Dumont Creek*

Although the eight observers' estimates (corrected) of reach surface area differed from the actual measured area, the differences were small. The 95% confidence intervals of all eight observers' estimates of reach area included the actual measured area of the reach (Table 2).

Individual observers identified 15–21 habitat units within the reach (mean, 19). The observers differed considerably in the amount of total reach area they classified as pools, riffles, and glides.

Differences among observers in classifying glides introduced substantial variability in habitat classifications. From 0 to 22% of the reach area was classified as glides. This discrepancy developed because habitat units with shallow (<0.5 m), slowly flowing water were often classified by some observers as pools and by others as glides. The proportion of the total area classified as riffles varied (from 31 to 44% of reach area), but classification disparity was less than for pools or glides (Table 2). The coefficients of variation (CV = 100 × SD/mean) were 118 for glides, 18 for pools, and 11 for riffles.

Although variation in classification of the primary habitat types was large, classification of

TABLE 3.—Consistency (% agreement) with which eight independent observers in trial 1 classified habitat units as the same pool type or as a glide. A percentage of 100 indicates unanimity among observers, whereas smaller percentages indicate inconsistency among observers.

| | Pool type | | | | | |
|---|---|---|---|---|---|---|
| Unit | Back-water | Scour | Dam | Trench | Plunge | Glide |
| 1 | 37.5 | 25.0 | 12.5 | 0.0 | 0.0 | 25.0 |
| 2 | 25.0 | 37.5 | 12.5 | 0.0 | 0.0 | 25.0 |
| 3 | 37.5 | 25.0 | 25.0 | 12.5 | 0.0 | 0.0 |
| 4 | 25.0 | 62.5 | 12.5 | 0.0 | 0.0 | 0.0 |
| 5 | 37.5 | 25.0 | 12.5 | 25.0 | 0.0 | 0.0 |
| 6 | 0.0 | 0.0 | 25.0 | 0.0 | 75.0 | 0.0 |
| 7 | 0.0 | 0.0 | 0.0 | 0.0 | 100.0 | 0.0 |
| 8 | 37.5 | 12.5 | 12.5 | 0.0 | 0.0 | 37.5 |
| 9 | 50.0 | 12.5 | 12.5 | 12.5 | 0.0 | 12.5 |

pools into secondary habitat types was even more variable (Table 3). In only one of nine cases did all observers agree on a specific pool type for a habitat unit that the majority of the observers had classified as a pool. In the other eight cases, observers differed in their classification of the pool type, which resulted in the same habitat unit being classified by observers as two to four pool types or as a glide.

### *Trials 2 and 3: Slick Creek and Lonewoman Creek*

The uniformly trained observers in trials 2 and 3 differed less in their classification of primary habitat types than did the less uniformly trained observers in trial 1. In the reaches of Slick and Lonewoman creeks, 53–60% of the combined 150 habitat units were classified as pools (Table 4). The 7% maximum difference among observers in the total number of pools was considerably less than the maximum difference of 32% in trial 1. Low variability among observers also existed in classifying riffles (36–49%) and glides (0–4%). The CVs in classification of habitat units into types were 5 for pools, 8 for riffles, and 44 for glides. As in trial 1, the observers differed most in their classification of glides.

Variability among observers in classifying pools, riffles, and glides was similar in trials 2 and 3. Within trial 2, the CVs were 8 for pools, 11 for riffles, and 64 for glides, and within trial 3, they were 7 for pools, 9 for riffles, and 73 for glides.

Coefficients of variation in the classification of secondary habitat types were considerably greater. The CVs were 13 for low-gradient riffles, 27 for high-gradient riffles, 43 for plunge pools, 46 for backwater pools, 55 for dam pools, 63 for lateral scour pools, 71 for trench pools, and 94 for cascading riffles.

TABLE 4.—Percentages of the combined 150 habitat units in trials 2 (Slick Creek) and 3 (Lonewoman Creek) classified by six observers as pool, riffle, and glide.

| | Percentage of area classified as | | |
|---|---|---|---|
| Observer | Pool | Riffle | Glide |
| 1 | 53 | 43 | 4 |
| 2 | 54 | 46 | <1 |
| 3 | 57 | 49 | 4 |
| 4 | 59 | 38 | 3 |
| 5 | 58 | 39 | 3 |
| 6 | 60 | 36 | 4 |

All six observers agreed on the primary habitat type of 110 of the combined 150 habitat units (73%). In trial 2, there was complete agreement among observers in classifying pools, riffles, or glides in 51 of the 75 (68%) habitat units; in trial 3, complete agreement was achieved in 59 of the 75 habitat units (79%). These differences were not statistically different ($\chi^2$ test, df 1, $P > 0.1$). Complete agreement on the secondary habitat type was significantly less common than complete agreement in the primary habitat types ($\chi^2$ test, df 1, $P < 0.01$) and occurred in only 34 of the 150 (23%) habitat units. Complete agreement among observers occurred in only 17 of 75 habitat units in both trial 2 and trial 3.

Observers were more consistent in classifying pool types in trial 3 (9 of 46) than in trial 2 (2 of 38). The greater amount of large woody debris in Lonewoman Creek (trial 3) may have accounted for the more consistent classification of pool types there. The nine pools in Lonewoman Creek on which all observers agreed were classified as plunge pools and 6 of 9 of these plunge pools were associated with large pieces of wood in the stream channel.

## Discussion

The variation among observers in the classification of primary and secondary habitat types appeared related to at least three factors: (1) the level of definition required in classification (e.g., pools in general versus specific types of pools), (2) the level and uniformity of observer training, and (3) the stream channel characteristics.

Generally, as the number of habitat types used to classify a stream increased, consistency among the observers decreased. For the combined 150 habitat units of trials 2 and 3, the observers were in complete agreement nearly 75% of the time when only three primary habitat types were involved, but they were in complete agreement less than 25% of the time when nine secondary habitat types were involved.

The inability of the observers to consistently distinguish among secondary habitat types appeared to be more acute for pools than for riffles, but results varied with stream reach. For example, in trial 2 observers demonstrated complete agreement in classifying secondary riffle types in 45% of the cases and secondary pool types in only 5% of the cases. In contrast, observers in trial 3 demonstrated complete agreement in classifying riffle types in 35% of the cases and pool types in 20% of the cases. The larger number of pool types compared to the number of riffle types (five secondary pool types versus three secondary riffle types) may have contributed to lower consistency in classifying secondary pool types.

Although our results indicated that consistency among observers improved with additional and uniform training, the use of an elaborate habitat classification scheme reduced repeatability despite the training provided. Our results indicated that 5 d of standardized instruction were insufficient to produce consistent habitat evaluations when a habitat classification scheme with nine habitat types was used.

Certain physical characteristics of streams may also influence the consistency with which habitat types are classified. For example, complete agreement among observers in classifying riffles in trials 2 and 3 may have been affected by the gradients of the two stream segments. Observers were less successful in consistently classifying riffles in trial 3 where the reach gradient was 4%, which is the accepted value for separating low-gradient from high-gradient riffles. Consequently, many habitat units in trial 3 likely had characteristics of both high-gradient and low-gradient riffles and thus were more difficult to classify than riffles in trial 2 where the stream gradient was 2.5%.

Although not tested in this study, discharge may also affect consistency of stream classification into habitat types. Dambacher (1991) reported that pool : riffle : glide ratios differed between 2 years in several Oregon streams. The greatest between-year variation was found in a stream where estimated stream volume more than doubled between years. Under the higher volume (and concurrent higher discharge and velocity) the tendency might be to classify more pools as riffles because areas of slow water would appear to be less common.

Our results from these three trials indicate that complex stream habitat classifications are not consistent among observers. It follows that, if stream surveys are conducted with complex classification systems (see McCain et al. 1991), the resulting descriptions of stream conditions are suspect because conclusions may depend as much on the way individual observers classified the stream channel as on the actual physical characteristics of the stream. We suggest that in many cases it may be more informative to conduct fewer, more rigorous studies describing characteristics of specific stream segments than to use complex habitat classification schemes to describe entire streams.

## Acknowledgments

We express special thanks to J. Dose, M. Faurot, and S. Platte for technical and logistical help in the collection of data for this study. T. C. Bjornn, D. H. Bennett, P. A. Bisson, J. L. Congleton, and G. H. Reeves critiqued drafts of the manuscript. Funding for this project was provided by the U.S. Forest Service, Umpqua National Forest, Roseburg, Oregon.

## References

Bisson, P. A., J. L. Nielsen, R. A. Palmason, and L. E. Grove. 1982. A system of naming habitat in small streams, with examples of habitat utilization by salmonids during low stream flow. Pages 62–73 *in* N. B. Armantrout, editor. Acquisition and utilization of aquatic habitat inventory information. American Fisheries Society Western Division, Bethesda, Maryland.

Dambacher, J. M. 1991. Distribution, abundance, and emigration of juvenile steelhead, and analysis of stream habitat in the Steamboat Creek basin, Oregon. Master's thesis. Oregon State University, Corvallis.

Hankin, D. G. 1984. Multistage sampling in fisheries research: applications in small streams. Canadian Journal of Fisheries and Aquatic Sciences 41:1575–1591.

Hankin, D. G., and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834–844.

Hawkins, C. P., and ten coauthors. 1993. A hierarchical approach to classifying stream habitat features. Fisheries (Bethesda) 18(6):3–12.

Heifetz, J., M. L. Murphy, and K. V. Koski. 1986. Effects of logging on winter habitat of juvenile salmonids in Alaskan streams. North American Journal of Fisheries Management 6:52–58.

Hicks, B. J. 1989. The influence of geology and timber harvest on channel morphology and salmonid populations in Oregon Coast Range streams. Doctoral dissertation. Oregon State University, Corvallis.

Knighton, D. 1984. Fluvial forms and processes. Edward Arnold, London.

Leopold, L. B., M. G. Wolman, and J. P. Miller. 1964. Fluvial processes in geomorphology. Freeman, San Francisco.

McCain, M., D. Fuller, L. Decker, and K. Overton. 1991. Stream habitat classification and inventory procedures for northern California. U.S. Forest Service, Fish Habitat Relationships, Technical Bulletin 1, Arcata, California.

Montgomery, D. R., and J. M. Buffington. 1993. Channel classification, prediction of channel response, and assessment of channel condition. Washington State Timber/Fish/Wildlife Agreement, TFW-SH10-93-002, Seattle.

Southwood, T. R. E. 1980. Ecological methods. Chapman and Hall, New York.

USFS (U.S. Forest Service). 1992. Stream survey handbook. USFS, Region 6, Portland, Oregon.